Structural Comparisons of the E. coli Prodrug-Activating Nitroreductase Enzymes, NfsA and NfsB

By

MARTIN ALAN DAY

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Abstract

The major and minor nitroreductase enzymes of *Escherichia coli*, NfsA and NfsB, have both been identified as potential prodrug activating enzymes in Gene-Directed Enzyme Prodrug Therapy (GDEPT) of cancer with 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB1954). NfsB mutants have been selected for their ability to sensitize cells to successively lower concentrations of CB1954. In this study the crystal structures of most improved double (T41L/N71S), triple (T41Q/N71S/F124T) and quadruple (T41Q/N71S/F124T/M127V) mutants have been determined using X-ray crystallography. The structures help explain the improvements seen with the mutants, in particular, the structure of the triple mutant reveals a hydrogen bond between Gln41 and Thr124, explaining why this mutation has greater activity than expected from the single mutants. Crystal structures of NfsB mutants where Phe124 is replaced with either of the unnatural amino acids *para*-nitrophenylalanine or *para*-aminophenylalanine show a shift in the side chain rotamer allowing these polar groups to interact with substrates, suggesting this leads to their increased activity. For the first time, the crystal structure of NfsA is presented bound to a series of ligands, including the inhibitors succinate and fumarate, and the substrate nitrofurantoin. Another structure bound to a second FMN shows how the flexible loop may move to form a phosphate binding pocket for NADPH.

For Tom, for making me smile and showing that there is more to life than everyone assumes. I will always miss you.

"There are more things in heaven and earth, Horatio, than are dreamt of in your Philosophy"

Hamlet Act 1, Scene 5, William Shakespeare

Acknowledgements

First and foremost I have to acknowledge and thank Eva for her constant help, guidance and support. I feel privileged to have had you for a supervisor. I would like to thank Scott for teaching and encouraging me, and for sparking my interest in crystallography, and Peter for his advice and encouragement. I would also like to thank Rosemary, mostly for her support and for always being there to talk to, and for helping me with anything in the lab that I was having trouble with. Thank yous also go to Andy and Klaus for some interesting discussions and advice, and to Nick and Baz for help with the stopped-flow in Birmingham. Thanks to all the people who have made the 7th floor such a great environment to work in, to Rich for going through it all with me, Rachel for showing me it's possible, Raul for making me laugh and smile, Mohammed for interesting discussions, and to Yu Pin, Sandeep, Craig, Jack and Matt for helping to keep me sane. Thanks go to the various undergraduate and masters students that have worked alongside me, in particular those whose work has contributed directly to this thesis; Alex Gavin, Hana Gannan, Asha Hassan, Sophie Englander and Laura Goodman. Also I would like to thank Simon Vass and Liz Hodgkins from Cancer Sciences, and Dr Ross Anderson, Dr Paul Race and Craig Armstrong at the University of Bristol for help with stopped-flow and advice. The opportunity to work at ESRF and Diamond was not only important to my project but great fun, so I would like to thank them and all the people there who gave me assistance. Finally I would like to thank my family for supporting me and making fun of me, and the Medical Research Council for funding me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations Used

Abbreviation	Meaning
ΔG	Gibb's free energy
3D	Three dimensional
ADEPT	Antibody-Directed Enzyme Prodrug Therapy
Ala	Alanine
Arg	Arginine
Asn	Asparagine
ASU	Asymmetric Unit
BSA	Bovine Serum Albumin
C_α	Carbon alpha of a amino acid
CB1954	5-(aziridin-1-yl)-2,4-dinitrobenzamide
CDEPT	Clostridial-Directed Enzyme Prodrug Therapy
CMDA	4-[(2-chloroethyl)(2-mesyloxyethyl)amino]benzyl-L-glutamic acid
DCIP	2,6-dichlorophenolindophenol
DEPT	Directed Enzyme Prodrug Therapy
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. cloacae	*Enterobacter cloacae*
E. coli	*Escherichia coli*
E_a	Thermodynamic free energy of activation
EDTA	Ethylenediaminetetraacetic acid
ESRF	European Synchrotron Radiation Facility, Grenoble, France
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
FRase I	Major NAD(P)H:FMN oxidoreductase from *Vibrio fischeri*
FRP	NADPH:flavin oxidoreductase from *Vibrio harveyi*
GDEPT	Gene-Directed Enzyme Prodrug Therapy
Gln	Glutamine
HIV	Human Immunodeficiency Virus
HPLC	High Performance (or Pressure) Liquid Chromatography
IC_{50}	Deffine!!!!!!
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k_{cat}	The Catalytic Rate Constant of an enzyme catalysed reaction
K_m	The Michaelis Constant of an enzyme catalysed reaction
K_{mapp}	The apparent Michaelis Constant
LB	Luria broth
Lys	Lysine
MDEPT	Macromolecule-Directed Enzyme Prodrug Therapy
mRNA	Messenger RNA
NAD	Nicotinamide adenine dinucleotide
NAD(P)H	Reduced form of NAD(P)
NADP	Nicotinamide adenine dinucleotide phosphate
Nap	Naphthylalanine

NAT2	Human N-Acetyltransferase II
NFN	Nitrofurantoin
NfsA	Major oxygen-insensitive nitroreductase from *Escherichia coli*
NfsB	Minor oxygen-insensitive nitroreductase from *Escherichia coli*
NFZ	Nitrofurazone
NMP	N-Methylpyrrolidone
NMR	Nuclear Magnetic Resonance
NOX	NADH oxidase from *Thermus thermophilus*
NR	Nitroreductase from *Enterobacter cloacae*
NRH	Dihydronicotinamide Riboside
NTR	Minor oxygen-insensitive nitroreductase from *Escherichia coli* (NfsB)
NQO1	NAD(P)H:quinone oxidoreductase 1or DT-Diaphorase
NQO2	NRH:quinone oxidoreductase 2
PAE	Prodrug-activating enzyme
pAF	*Para*-aminophenylalanine
PCR	Polymerase Chain Reaction
PDEPT	Polymer-Directed Enzyme Prodrug Therapy
PEG	Polyethylene Glycol
Phe	Phenylalanine
PMSF	Phenylmethanesulphonylfluoride
pNF	*Para*-nitrophenylalanine
PSA	Prostate-Specific Antigen
Quinol	Hydroquinone or Benzene 1,4-diol
Quinone	1,4-Benzoquinone
rEK	Recombinant Enterokinase
RNA	Ribonucleic Acid
S. typhimurium	*Salmonella typhimurium*
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SER	Surface Entropy Reduction
SKOV-3	Human ovarian carcinoma cell line
STE	Salt Tris EDTA buffer
Thr	Threonine
TNT	Trinitrotoluene
tRNA	Transfer RNA
Tyr	Tyrosine
Val	Valine
VDEPT	Virus-Directed Enzyme Prodrug Therapy
WT	Wild-Type
YfkO	Nitroreductase from *Bacillus licheniformis*
YieF	Chromate reductase from *Escherichia coli*

1. Introduction

1.1 Enzymes as Biological catalysts

1.1.1 The Nature of Catalysis

Many chemical reactions would not be likely to occur without catalysts. The rate of any chemical reaction is limited by the free-energy of activation (E_a), and catalysts work by lowering this (Figure 1-1). Catalysts provide a different transition state with a lower E_a (as shown in red in Figure 1-1) and thereby allow reactions that would normally be unlikely to occur because some kinetic or energetic barrier. However it does not change the energy difference between the starting reactants and the final products, called the Gibbs free energy or ΔG. This lower E_a allows reactions to occur at a faster rate than would be possible without a catalyst at the same temperature, as more reactants will have the required energy to react. The E_a is related the the rate constant (k) for a chemical reaction through the Arrhenius equation (equation 1.1). Where T is temperature in Kelvin, R is the universal gas constant and A is the pre-exponential factor. It is the nature of catalysts that they are not themselves altered at the end of the reaction and do not alter its extent or equilibrium. They can be altered during the reaction but are always regenerated by the end.

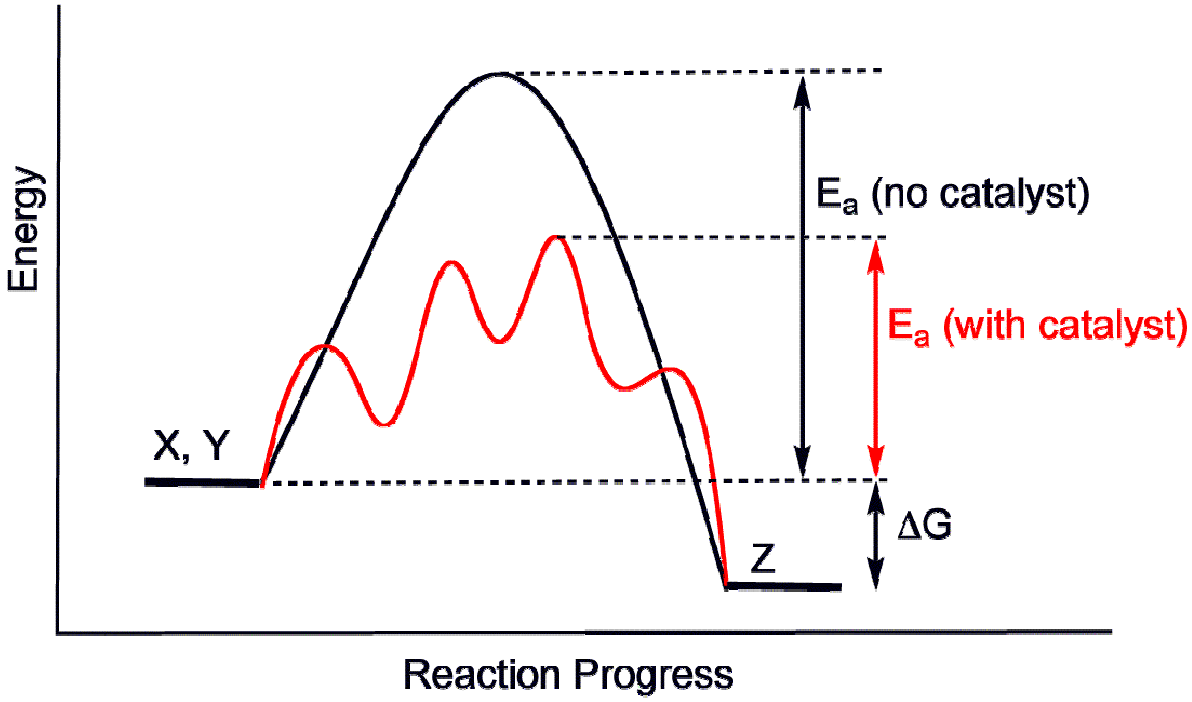


Figure 1-1: The Thermodynamic effect of catalysis on E_a. A reaction of X and Y to give Z (with a Gibb's free energy, ΔG, the energy difference between the two states) is shown, with the effect of the catalyst on the E_a shown in red, where a new pathway of reaction is opened up. The peaks represent transition states and the troughs reactive intermediates. Figure from http://en.wikipedia.org/wiki/Catalysis.

Equation 1-1

$$k = Ae^{-E_a/RT}$$

1.1.2 Discovery and early understanding of enzymes

Enzymes are the catalysts of the biological world, and make up the majority of proteins found in bacterial and mammalian cells. The term enzyme was first coined by Wilhelm Kühne in 1878 and comes from the Greek 'in leaven' (Kühne, 1877), from the discovery that the catalytic activity of cells came from within not the whole microrganism. At this point humanity had been utilizing enzymes for centuries

without knowing it in processes such as fermentation and cheese-making. Eduard Buchnar in 1897 showed for the first time isolated enzymatic activity in the absence of whole cells, in discovering that yeast extracts alone could still catalyze fermentation, he received the Nobel Prize for his work in 1907 (Buchner, 1897). In 1894 Emil Fischer proposed the 'lock and key' hypothesis of enzyme specificity (Figure 1-2) based upon experiments performed with carbohydrate metabolising enzymes, where the enzyme's specificity for its substrate was shown (Fischer, 1894).

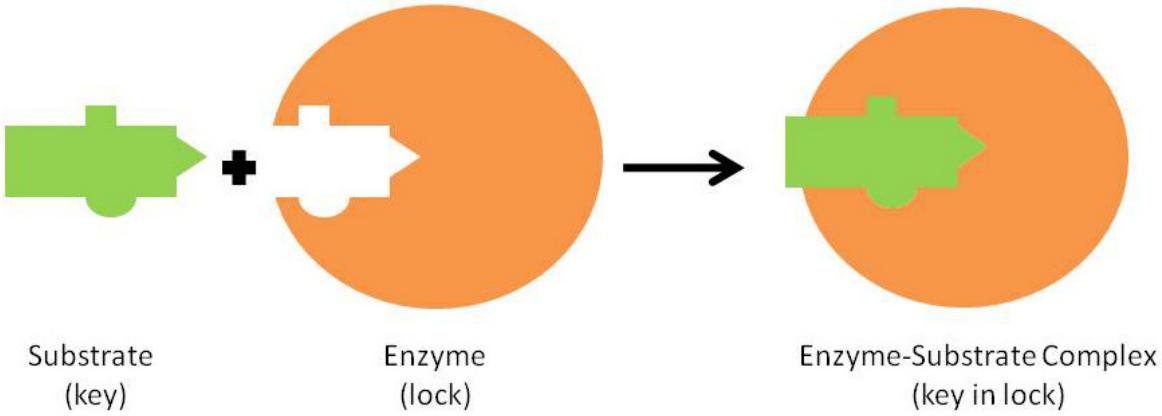


Figure 1-2: The lock and key hypothesis of enzyme action. The substrate (green) and enzyme (orange) are shaped such that the substrate has to be approximately the right shape for the enzyme.

The chemical makeup of enzymes was unknown until 1926 when Sumner crystallized urease, which catalyses the hydrolysis of urea to carbon dioxide and water, and it was shown that it was made up entirely of amino acids linked through amide bonds, i.e. protein (Sumner, 1926). This chemical nature of catalysis may have been missed if Sumner's techniques had been more sensitive, as they would have revealed a small amount of Nickel which is essential for catalysis. Svedberg's development of the

ultracentrifuge (also in the 1920's, for which he won the Nobel Prize in 1927), allowed wide-ranging isolation of enzymes and showed (along with the earlier crystallization) that they were individually isolatable molecules of specific masses (generally with M_r's ranging from 10^4 to 10^7 Da), instead of colloidal suspensions. This suggested that enzymes were not random polymers but organised structures whose chemical action could be deduced, and this was confirmed in the 1960's, firstly with the determination of the amino acid sequence of ribonuclease (Hirs *et al.*, 1960), and secondly with the determination of the three-dimensional structure of lysozyme by X-ray crystallography (Blake *et al.,* 1965). Today, the structures of many thousands of enzymes (and other proteins) have been determined.

Studies in the 1950's and 1960's suggested that enzymes showed greater flexibility than was implied by the 'lock and key' hypothesis, and in 1958 the induced fit model (Figure 1-3) was proposed by Koshland to account for the catalytic power and specificity of enzymes (Koshland, 1958). It became obvious that enzyme activity could be controlled by chemical conditions in the cell and the 'allosteric model' of enzyme action was proposed by Monod (Monod *et al.,* 1963, Monod *et al.*, 1965). This model described how small chemical molecules, called 'effectors', could alter the activity of an enzyme. Effectors called 'activators' enhance an enzyme's activity and allosteric 'inhibitors' decrease its activity. Effectors may bind away from the active site and it has been postulated that part of thier action may be through the induction of structural changes in the enzyme upon effector binding.

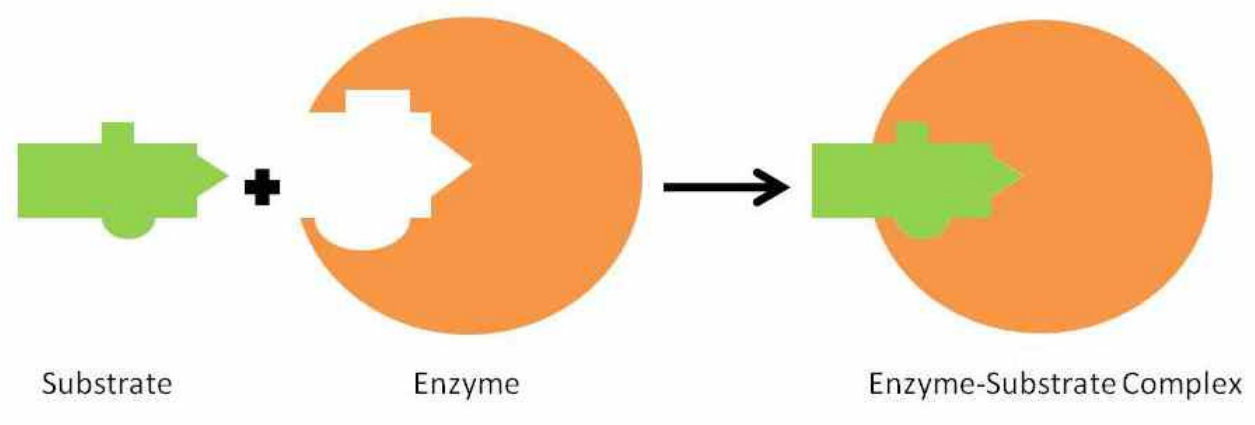


Figure 1-3: Induced fit hypothesis of enzyme specificty, showing how the enzyme (orange), may change shape upon binding of the substrate (green).

Enzymes are distinguished from non-biological catalysts primarily by their high catalytic activity (up to 10^{17} fold improvement in rate of reaction), their specificity and the fact that they can be regulated by chemicals. Enzymes may work by lowering the energy (E_a) of a transition state, providing a route to a new transition state with a lower E_a or by bringing the reactants together into the correct orientation, effectively an entropic effect not affecting the enthalpy.

1.2 Flavoenzymes

The focus of this thesis is the nitroreductase enzymes of *Escherichia coli*. These belong to a family of enzymes called flavoenzymes, a family of yellow coloured enzymes that catalyze a wide range of biological processes ranging from energy production to biosynthesis (reviewed in (Joosten and van Berkel, 2007)). They catalyze a diverse range of one- or two-electron reductions, giving them a key role in metabolism. Flavoenzymes are characterised by their tightly bound flavin cofactors,

either flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). The flavin can also be either non-covalently (e.g. NfsA and NfsB, discussed later) or covalently anchored to histidine, cysteine or tyrosine residues (e.g. vanillyl-alcohol oxidase (Fraaije *et al.*, 1999)), and more recently it has been discovered that this can also be via a dual interaction with two residues (Huang *et al.*, 2005, Winkler *et al.*, 2006, Rand *et al.*, 2006). The unique spectroscopic properties of the flavin cofactors make flavoenzymes perfect for enzymological study (Massey, 1995), and they are thus one of the best characterised families of enzymes.

The 3D structures of many flavoenzymes have been determined. In fact to give an indication of the extent of study of flavoenzymes, at present the protein data bank (www.pdb.org) contains 1,272 structures containing FAD and 590 structures containing FMN (out of a total of 85,212 structures, as of 14/10/2012). There is a variety of different structural topologies seen across these structures, with enzymes performing similar functions sometimes sharing and sometimes differing in their secondary structure motifs (Fraaije and Mattevi, 2000).

1.2.1 Flavin-binding

The flavin cofactors are the site of catalysis in flavoenzymes. Figure 1-4 a) shows the structure of the flavins FMN and FAD. Both contain an identical tricyclic isoalloxazine ring, an amphipathic molecule combining a hydrophobic dimethylbenzene moiety and a hydrophilic pyrimidine ring. The isoalloxazine ring does slightly deviate from planarity, although the degree of this deviation does not

seem to affect the degree of hydrogen bonding to the peptide (Fraaije and Mattevi, 2000). The ability of flavins (and flavoenzymes) to catalyse redox reactions is due to their ability to adopt three different redox states; reduced, oxidised and semiquinone (Figure 1-4b) (Mayhew, 1999). It is the flavin's ability to form the semiquinone that allows flavoenzymes to catalyse both one- and two-electron reactions. The two-electron reduction of free flavin has a redox potential of around -200 mV, but upon binding to enzymes this can change greatly, ranging from approximately +60 mV to -400 mV (Ghisla and Massey, 1989, Fraaije *et al.*, 1999, Mayhew, 1999). Areas of positive charge around isoalloxazine rings are thought to increase the redox potential, whilst areas of negative charge or hydrophobicity lower it (Massey, 1995, Ghisla and Massey, 1989). FAD has been shown to bind covalently in some cases, and site-directed mutagenesis studies have suggested that covalent interactions could increase the oxidative power of the flavin (Fraaije *et al.*, 1999). These alterations to the flavin properties, caused by their binding to enzymes, increase the range of reactions that flavoenzymes can catalyse and also show the importance of the enzyme itself in the reactions at the flavin.

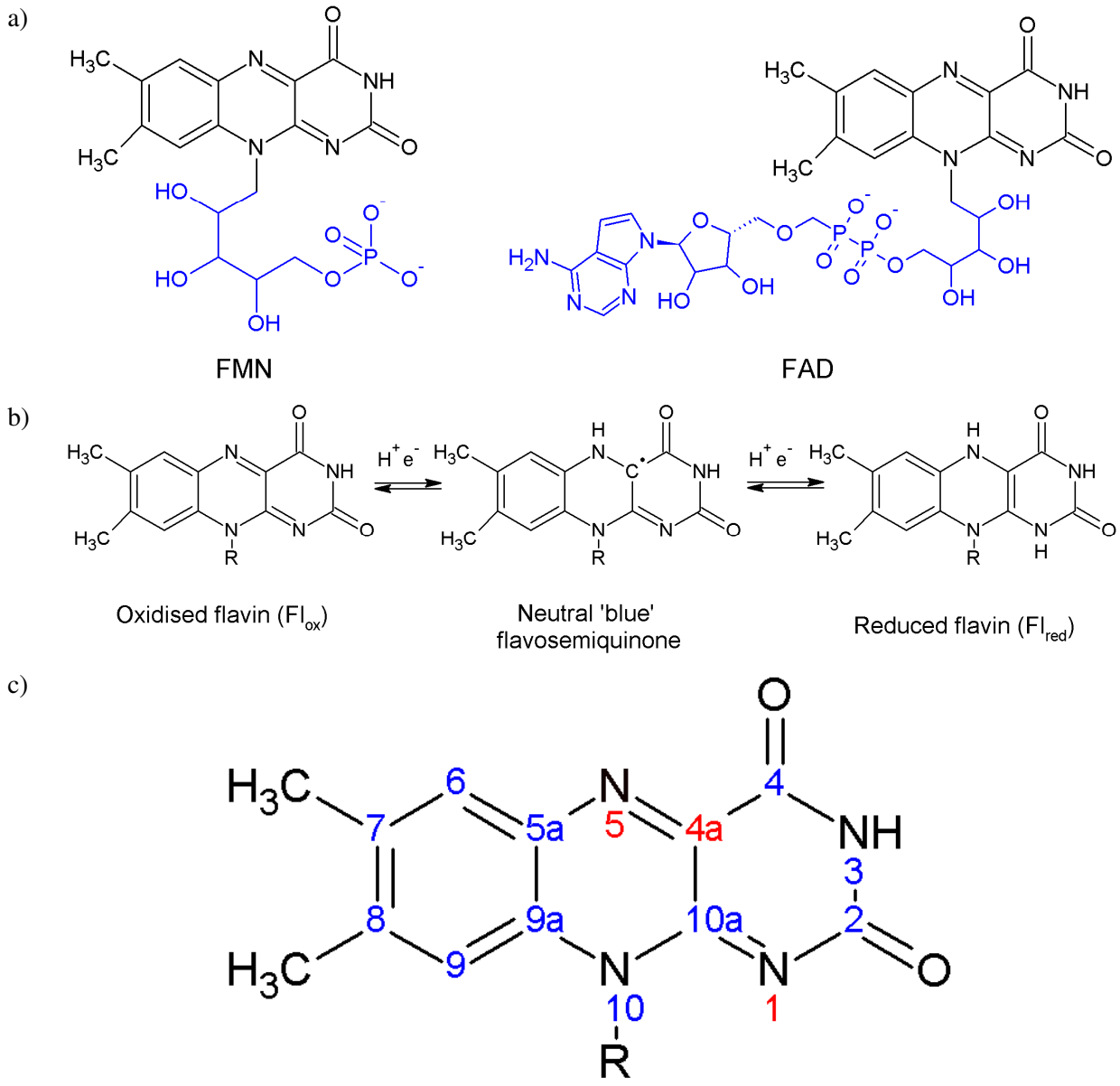


Figure 1-4: a) The chemical structure of FMN and FAD. The isoalloxazine ring is shown in black. b) The various oxidations states of the isoalloxazine ring of the flavin. c) The numbering of the atoms of the isoalloxazine ring. The isoalloxazine ring has two faces, *re* and *si*, and in all images the *si* side is facing the viewer (Mayhew, 1999). Figure produced using ACD/ChemSketch Freeware, version 10.00, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2012.

The possible oxidation states of the flavins are shown in Figure 1-4 b). The reduced form is often deprotonated at the N1 (or N5) nitrogen (see Figure 1-4 c) for numbering

of the isoalloxazine rings), and this delocalised anionic state at the N1-C2=O2 locus is usually stabilized through contact (< 3.5 Å distance) with positively charged parts of the protein (Fraaije and Mattevi, 2000). This may be due to side chains such as lysine or arginine, which are fully charged, or to the N-terminus of an α-helix or a group of peptide nitrogens, which are polar. These residues are key to FMN binding and increase the redox potential of the cofactor, in dihydroorotate dehydrogenase (Bjornberg *et al.*, 1997) and trimethylamine dehydrogenase (Mewies *et al.*, 1996), replacing the positively charged sidechain stabilising the N1-C2=O2 locus with a neutral one caused the enzyme to become inactive and prevented FMN binding. There are multiple types of flavoenzyme, which are discussed in more detail elsewhere (Massey, 1995).

1.2.2 Reduction of the flavin

Many flavoenzymes catalyse dehydrogenation reactions, by which a substrate is oxidised and the flavin is reduced (at the N5 nitrogen). This reactive centre is usually within hydrogen bonding distance of a hydrogen bond donor from the protein itself (Fraaije and Mattevi, 2000). This interaction also has a highly conserved stereochemistry, forming an angle of between 116° and 170° from the N10-N5 to the donor residue (green in Figure 1-5). The N5 nitrogen of the isoalloxazine ring is one of two which are protonated in the reduction reaction displayed in Figure 1-4b and, unlike the N1, is unlikely to be deprotonated in solution. When crystal structures of flavoenzyme ligand complexes are studied, the N5 shows a highly conserved interaction with the ligand at a distance of around 3.5 Å and at an angle of between 96° and 117° from N5-N10 (red in Figure 1-5) (Fraaije and Mattevi, 2000). So the

interactions with the N5 and the centre of the isoalloxazine ring system are important in binding and activity of the enzyme. The active part of the substrate often binds in a buried part of the structure and is thus protected from solvent.

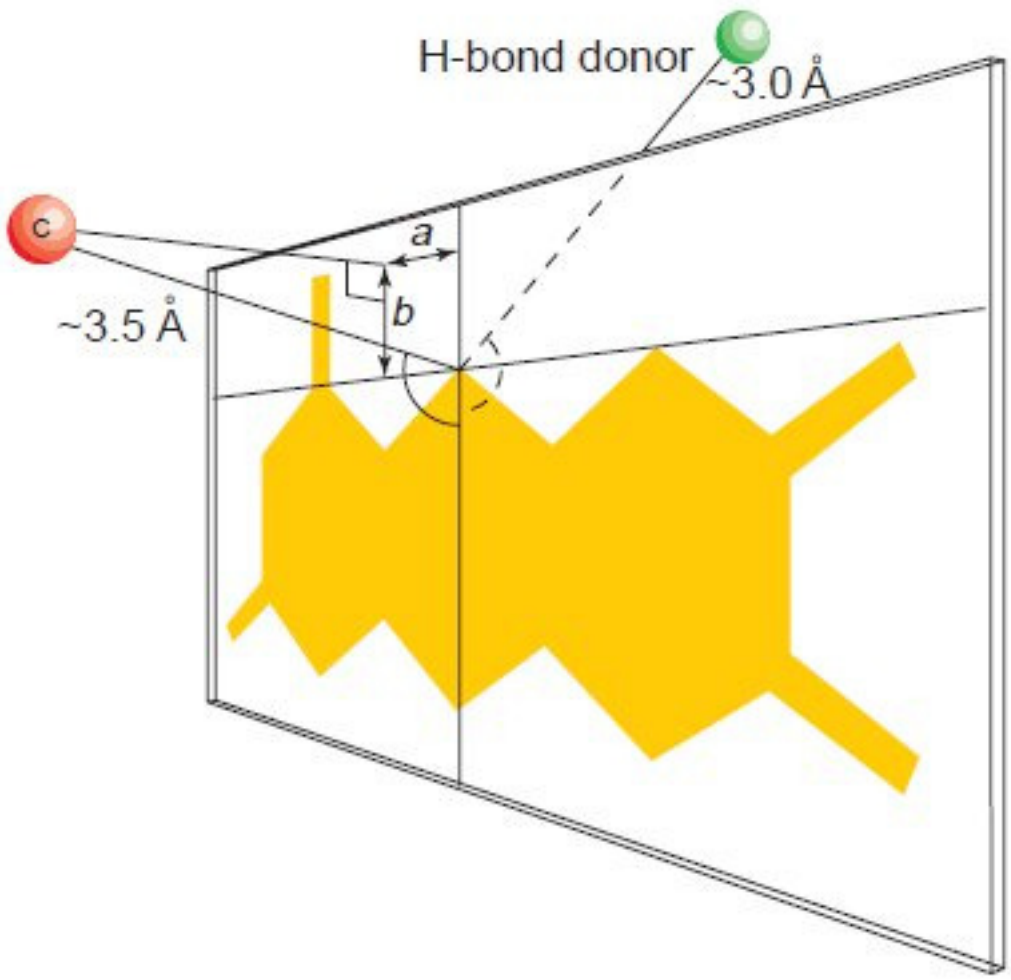


Figure 1-5: The distances between the reactive flavin N5 and the hydrogen-bond donor (green), and the distance from the flavin of the CH group that is oxidised (red – see the reaction mechanism in Figure 1-6). The distances of *a* range from 0.2 - 1.1 Å, and *b* from 0.5 – 1.4 Å. The distances are based on those of a number of crystal structures detailed further in the reference from which this figure is also adapted (Fraaije and Mattevi, 2000).

Several crystal structures of reduced flavoenzyme have been determined. These include the nitroreductase enzyme from *Enterobacter cloacae* (Haynes *et al.*, 2002), where the isoalloxazine ring of the reduced flavin is in a more strongly bent conformation than the oxidised flavin. The angle changes from 16° in the oxidised enzyme to 25°, which is similar to that expected for free reduced flavin in solution, as the addition of a proton favouring a bent conformation. Oxidised flavin free in solution would be expected to be planar, and thus the bent conformation is induced by the enzyme in order to favour reduction of the flavin. However, it is possible that the flavin could be photoreduced in the X-ray beam.

1.2.3 Mechanism of reactions

Flavins perform two-electron substrate-dehydrogenation reactions, in which two electrons are transferred from the substrate to the flavin and a hydrogen is lost from the substrate and gained by the flavin at the N5 nitrogen. There are three potential mechanisms for this reaction, one a direct transfer of a hydride, one a carbanion mechanism and the final one a radical mechanism. These are reviewed by Fraaije *et al.* (Fraaije and Mattevi, 2000). The three potential mechanisms are detailed in Figure 1-6. It is possible that all three exist and that flavoenzymes could also catalyse the reverse reactions.

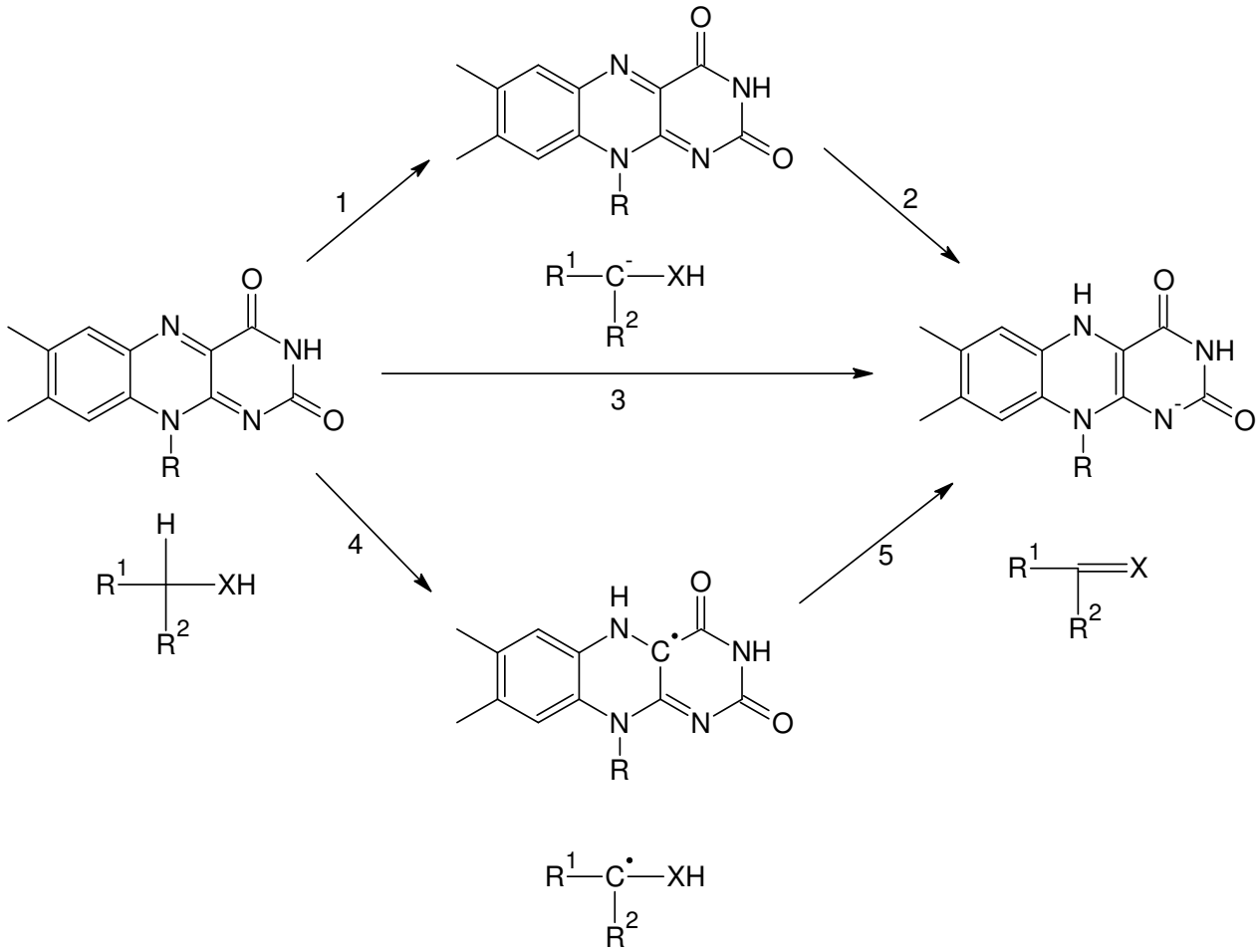


Figure 1-6: The different reaction mechanisms of flavin-catalysed substrate-dehydrogenation reactions – 1 and 2, the carbanion mechanism, 3, direct hydride transfer, and 4 and 5, the radical mechanism. R^1 and R^2 are non-specific groups attached to the reactive CH, whilst XH is the activating group. Figure adapted from reference (Fraaije and Mattevi, 2000) and produced using ChemSketch (ACD/labs).

1.3 Nitroreductase Enzymes

1.3.1 The bacterial nitroreductases

Nitroreductases are a family of enzymes that catalyze the reduction of many quinone and nitroaromatic molecules inside the cell. They are found in bacterial and mammalian systems. The bacterial nitroreductases are a family of flavoenzymes that catalyse the NAD(P)H dependent reduction of nitro groups on nitroaromatic and nitroheterocyclic compounds (Roldan *et al.*, 2008). This reduction can either be via a one-electron or a two-electron process; hence the nitroreductases can be grouped into two broad classes: Type I or oxygen-insensitive nitroreductases use two electrons from NAD(P)H (Figure 1-7) to catalyze the two-electron reduction of nitro groups to nitroso (and subsequently hydroxylamine and amino derivatives, although this may or may not be enzyme catalyzed) (Bryant *et al.*, 1981, Peterson *et al.*, 1979). The nitroso intermediate is usually not observed, as a second two-electron reaction reduces it rapidly to the hydroxylamine product, which may be the end point of the reaction. There are mammalian enzymes that also catalyze similar reactions, examples being the NAD(P)H:quinone reductase or DT-diaphorase (Knox *et al.*, 1993b) and xanthine oxidase (Ueda *et al.*, 2003). Type II or oxygen-sensitive nitroreductases use single electrons to catalyze the reduction of nitro groups to the nitro anion radical. Oxygen can catalyse the reverse reaction creating a futile cycle. A second one-electron reaction may occur to form the nitroso intermediate and harmful superoxide radicals can be formed. Enzymes that catalyze similar reactions are found in many mammalian systems (Mason and Holtzman, 1975a, Peterson *et al.*, 1979).

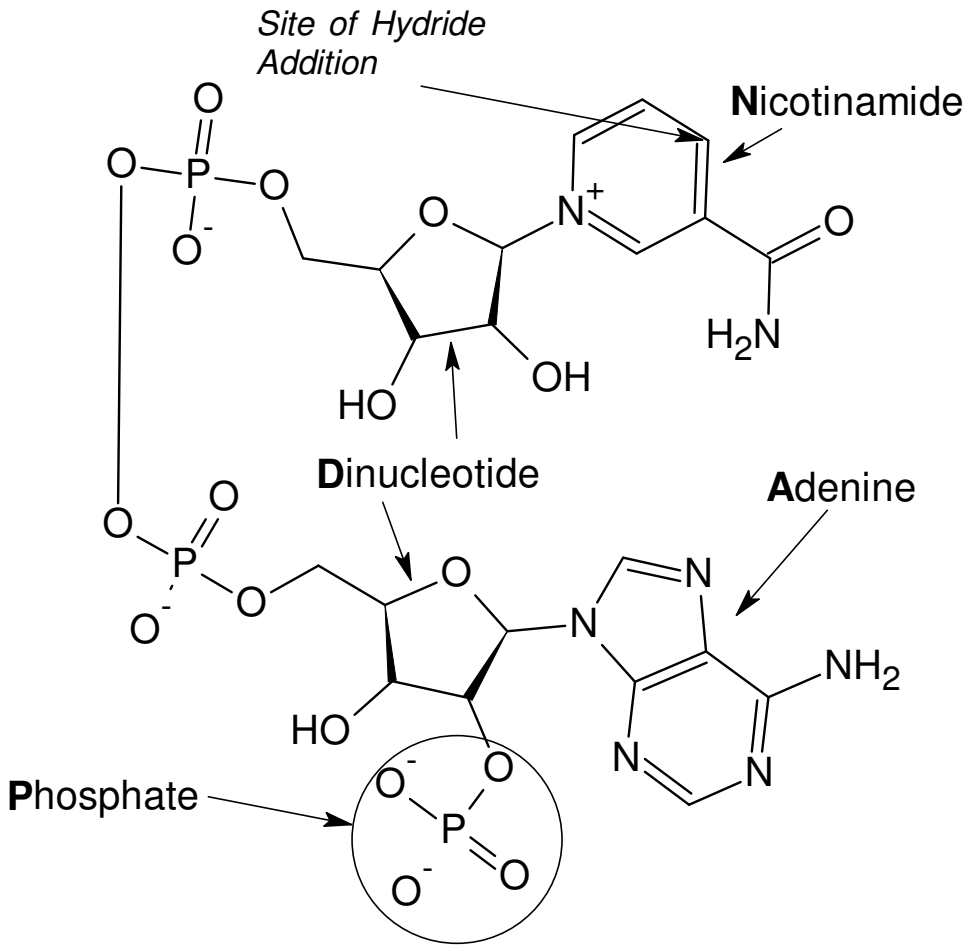


Figure 1-7: Chemical structure of $NADP^+$ showing each component. NADH lacks the 2' phosphate circled. NAD(P)H is reduced at the site indicated on the Nicotinamide ring. Figure produced using ChemSketch (ACD/labs).

The physiological role of the bacterial nitroreductases is still unknown, although it has been suggested that such enzymes may protect organisms from oxidative damage caused by semiquinones produced from one-electron reductions by other enzymes such as cytochrome P450 (Monks *et al.*, 1992). Several enzymes have seen industrial or clinical uses; the nitroreductase of *S. typhimurium* is used as part of the Ames mutagenicity test (Hajos and Winston, 1991), whilst the nitroreductases from *E. cloacae* and others have been identified as having potential use in bioremediation,

since they can reductively break down the explosive TNT (Kitts *et al.*, 2000, Spain, 1995, Esteve-Nunez *et al.*, 2001).

1.3.2 Discovery of the *E. coli* nitroreductases

The first evidence of a nitroaromatic reductase activity in *E. coli* came when the observation was made that whole cell preparations were capable of reducing the antibiotic chloramphenicol (Smith and Worrel, 1949). There were later shown to be two distinct types of nitroreductase present (Asnis *et al.*, 1957), which were shown to be type I and type II nitroreductases (McCalla *et al.*, 1970, Mason and Holtzman, 1975b, Peterson *et al.*, 1979). Investigation into the mechanism of resistance in *Escherichia coli* to nitrofurazone (NFZ) and other nitrofuran antibiotics led first to the discovery that nitrofurazone-resistant strains lacked a 'nitroreductase' activity (McCalla *et al.*, 1971), and later to the discovery of two 'nitrofuran sensitivity genes', *nfsA and nfsB* (McCalla *et al.*, 1978). The genes, later also referred to as *nfnA* and *nfnB*, were mapped to 79.8 min and 13 min respectively, following the isolation of mutants resistant to nitrofurantoin, and thus are under the control of different operons (Sastry and Jayaraman, 1984). Mutations in these genes provided resistance to nitrofuran antibiotics, including nitrofurazone and nitrofurantoin (Whiteway *et al.*, 1998). The proteins encoded by these genes are examples of bacterial nitroreductases.

1.3.3 Nitroreductase activity in *E. coli*

Of the type I and II nitroreductase activity in *E. coli,* the oxygen-insensitive nitroreductases show higher activity than the oxygen-sensitive with nitrofurazone as a

substrate. Further investigation into the oxygen-insensitive nitroreductase activity in *E. coli* concluded that it was the product of three genes and not just the two previously identified (Bryant *et al.*, 1981). The highest of the three nitroreductase activities with nitrofurazone was NADPH dependent; this is the product of the previously described *nfsA* gene (McCalla *et al.*, 1978). The other two activities could use either NADH or NADPH; one of these was the product of the previously described *nfsB* gene. The NfsA and NfsB proteins, the products of the *nfsA* and *nfsB* genes, are discussed in more detail later. Unlike *nfsA* and *nfsB*, the final gene conferring a type I nitroreductase activity had not been previously described and has not been genetically mapped. Little is known about the gene, named *nfsC*, or its product enzyme NfsC, other than that it appears to be rendered inactivate by $MnCl_2$, which had previously been used in a purification step (Bryant *et al.*, 1981).

E. coli also contains a minimal nitroreductase called ydjA, one of the smallest nitroreductase enzymes discovered to date (Choi *et al.*, 2008). YdjA was discovered when the genome of *E. coli* K12 (Blattner *et al.*, 1997) showed the existence of a putative nitroreductase of only 190 amino acids (most nitroreductases are >210) that was conserved across several organisms (Choi *et al.*, 2007). While it shares very little sequence homology with *nfsA* or *nfsB*, it is homologous to a putative nitroreductase from *Haemophilus influenza* (Fleischmann *et al.*, 1995), which belongs to a nitroreductase-like family that includes NADH oxidase and arsenite oxidase.

1.4 Gene therapy with *E. coli* Nitroreductase Enzymes

NfsA and NfsB both have been identified as having a potential for use in gene therapy of cancer with the prodrug CB1954 (Knox *et al.*, 1992, Vass *et al.*, 2009).

1.4.1 Principles of Suicide Gene Therapy

In cancer chemotherapy, drug doses are often limited by fatal side effects. Most chemotherapeutic drugs target all rapidly dividing cells, not just cancerous ones. Current research into cancer treatment is focussed on finding methods that will specifically target only tumour cells. Suicide gene therapy is a potentially viable, tumour-focused, technique. An exogenous enzyme is transferred (or produced) selectively within tumour cells. Then a non-toxic prodrug is administered, which is selectively activated to its cytotoxic product by the enzyme at the site of the tumour. The prodrug is converted to a cytotoxic derivative by the prodrug activating enzyme (PAE). It is important to use an enzyme not normally found in human cells, so the prodrug is only activated in the specific location containing the exogenous enzyme. This is illustrated in Figure 1-8 using the example of the nitroreductase/CB1954 combination discussed in Section 1.4.4.

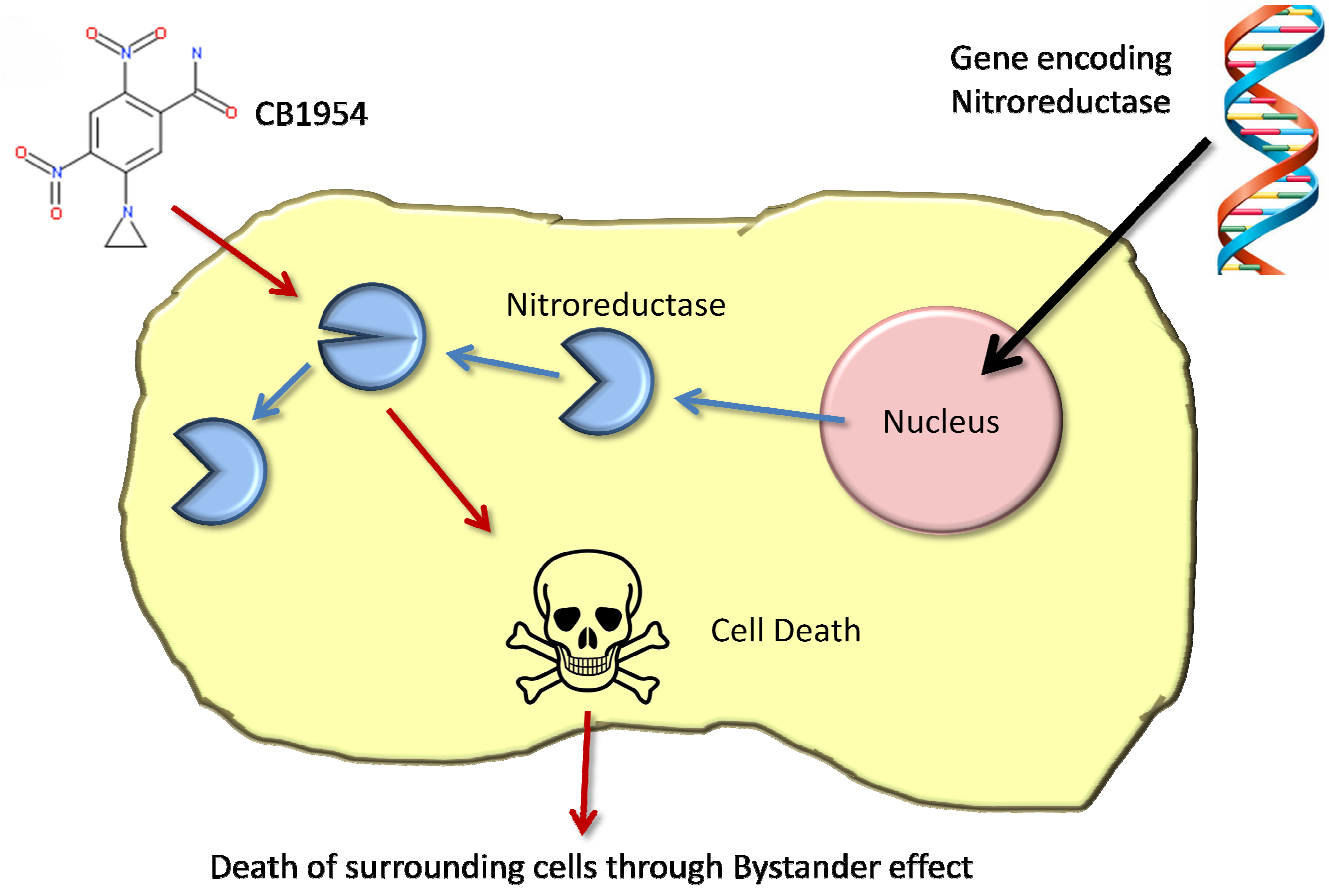


Figure 1-8: Illustration of the principles of GDEPT with nitroreductase and CB1954. The gene encoding the nitroreductase is inserted into a cell by means of an adenovirus vector. The gene is incorporated into the host cell and the nitroreductase is expressed. 48 hours later CB1954 is administered intravenously and only becomes active in cells expressing nitroreductase. This leads to the death of the cell and of neighbouring cells through a Bystander Effect (the local spread of the reactive product of CB1954 reduction).

A potential advantage to this type of treatment over conventional chemotherapy is that the local generation of cytotoxic derivatives gives lower levels of systemic toxicity (Kerr *et al.*, 1997, Niculescu-Duvaz and Springer, 1997, Russell and Khatri, 2006). Since it is unlikely that every cell would express the PAE, the ability of the cytotoxic drug to spread and cause toxicity in surrounding cells is vital to any form of suicide gene therapy. This is called the 'bystander effect' (Freeman *et al.*, 1993). The requirement for the spread of the cell-killing compounds is one of the problems associated with gene therapy. The spread ideally needs to avoid healthy tissue to reduce unwanted side effects. The major challenges of gene therapy remain to the method of getting the enzyme into the cell (both the vector for transport and the physical method of transporting the vector into the tumour), and the viability and efficiency of the enzyme prodrug combination. Even so, gene therapy is a promising avenue for cancer research and is showing some good early results (Tietze and Schmuck, 2011).

1.4.2 Types of Directed Enzyme Prodrug Therapies

A vector is required to target the enzyme to the tumour in order for the prodrug to be selectively activated at the site of the tumour. This selective targeting of an enzyme to a tumour is called Directed Enzyme Prodrug Therapy (DEPT). In Antibody Directed Enzyme Prodrug Therapy (ADEPT), the PAE is directed to the outside of target cells using a tumour specific antibody (Bagshawe, 1987, Bagshawe, 1990). A number of ADEPT combinations have been described that have shown promising results *in vitro* and *in vivo*, including carboxypeptidase G2 and para-N-bis (2-chloroethyl) aminobenzoylglutamic acid (Bagshawe *et al.*, 1988, Bignami *et al.*, 1992, Haisma *et*

al., 1992, Senter *et al.*, 1991, Roffler *et al.*, 1991). *E. coli* NfsB was proposed as a potential enzyme for use with ADEPT (Anlezark *et al.*, 1992, Anlezark *et al.*, 1995, Knox *et al.*, 1993b, Knox *et al.*, 1995), however as the enzyme is not internalised into the cell it does not have access to the first substrate (NAD(P)H). In order for the system to be viable alternative methods would be required that would lead to the enzyme being produced inside the cell. An alternative tumour directing method would be to use a polymer-based system similar to ADEPT called Macromolecule Directed Enzyme Prodrug Therapy (MDEPT) (Skelly *et al.*, 2001) or Polymer-Directed Enzyme Prodrug Therapy (PDEPT) (Satchi *et al.*, 2001). Several enzyme prodrug combinations have also been trialled in Clostridial-Directed Enzyme Prodrug Therapy (CDEPT), in which the Clostridia bacteria is used to transfer the enzyme to the tumour (Minton, 2003, Theys *et al.*, 2006). This targets the hypoxic areas of tumours by exploiting the anaerobic growth of Clostridia.

Gene Directed Enzyme Prodrug Therapy (GDEPT) is a gene therapy system that would allow the enzyme to be produced inside the tumour cells themselves. The gene encoding a non-human enzyme is inserted directly into the tumour cells, usually using either a virus or a plasmid, and the gene is expressed producing the enzyme. GDEPT has been trialled for use with several enzyme-prodrug combinations including cytosine-deaminase with 5-fluorocytosine (Brown and Lemoine, 2004), herpes simplex virus thymidine kinase with ganciclovir (Fillat *et al.*, 2003), and *E. coli* NfsB with the prodrug 5-(Aziridin-1-yl)-2,4-dinitrobenzamide (CB1954) (Denny, 2002, Patel *et al.*, 2009). A variation on GDEPT is Virus Directed Enzyme Prodrug Therapy (VDEPT), in which a viral vector is injected directly into the tumour to transfer the

gene encoding a non-human enzyme (Huber *et al.*, 1994). This system had been trialled for use with NADPH-cytochrome P450 reductase with cyclophosphamide (Tychopoulos *et al.*, 2005), and *E. coli* NfsB with CB1954 (Patel *et al.*, 2009).

1.4.3 Enzyme Prodrug Combinations

There have been several enzyme prodrug combinations tested, and a number of these are summarised in Table 1-1. Each of the enzymes activates the drug, but the drugs cause cell death in different ways. The combination of NfsB nitroreductase and CB1954 (as well as Cytochrome P-450 and cyclophosphamide) has the advantage over other DNA-alkylating agents that it can form DNA crosslinks, making it more potently cytotoxic as these are poorly repaired. This combination for gene therapy also produces a DNA crosslinking agent produced that has the ability to kill both dividing and non-dividing cells by inducing apoptosis (Drabek *et al.*, 1997, Knox *et al.*, 1988b). Other advantageous features include the fact that apoptosis occurs via a p53-independent mechanism (Cui *et al.*, 1999), the fact that only three cycles of prodrug administration were required in xenograft models showing the efficacy of this combination (McNeish *et al.*, 1998, Djeha *et al.*, 2000).

Enzyme	Prodrug	Method of Toxicity	Reference
Thymidine Kinase	Ganciclovir	Generation of Anti-Metabolites	(Moolten, 1986)
Cytosine Deaminase	5-Fluorocytosine	Generation of Anti-Metabolites	(Austin and Huber, 1993)
Deoxycytidine Kinase	Cytosine Arabinoside	Generation of Anti-Metabolites	(Hapke *et al.*, 1996)
Carboxypeptidase	CMDA (4-[(2-chloroethyl)(2-mesyloxyethyl)amino]benzyl-L-glutamic acid)	Production of DNA-Alkylating agents	(Marais *et al.*, 1996)
Cytochrome P-450	Cyclophosphamide	Production of DNA-Alkylating agents	(Chen *et al.*, 1996)
	2-Aminoanthracene	Production of DNA-Alkylating agents	(Rainov *et al.*, 1998)
	Paracetamol	Interaction with tissue macromolecules	(Thatcher *et al.*, 2000)
Linamarase	Linamarin	Production of cyanide	(Cortes *et al.*, 1998)
Nitroreductase	CB1954	Production of DNA-Alkylating agents	(Bridgewater *et al.*, 1995)

Table 1-1: A summary of various enzyme-prodrug combinations trialled for directed enzyme prodrug therapy of cancer, and the method by which each causes cell toxicity.

1.4.4 Development of Gene Therapy with Nitroreductases and CB1954

1.4.4.1 DT-Diaphorase and CB1954

CB1954 was discovered to potently inhibit the growth of Walker rat carcinoma 256, but not of other related tumours, which had previously been shown to be inhibited by DNA alkylating agents (Cobb *et al.*, 1969). It was shown that inside the Walker tumour cells *in vitro* (but not inside Chinese hamster cells) CB1954 was converted to a difunctional alkylating agent that caused DNA-DNA interstrand crosslinks (Roberts *et al.*, 1986). It was postulated that there was some activity within the Walker carcinoma cells that converted CB1954, and the search was on for a human tumour that would act in a similar way. CB1954 was described as "a drug in search of a human tumour to treat" (Workman *et al.*, 1986). The enzyme responsible for 'activating' CB1954 in Walker carcinoma cells was shown to be a 33.5 kDa enzyme called DT-Diaphorase or NAD(P)H:quinone oxidoreductase 1(NQO1) (Knox *et al.*, 1988b).

NQO1 is a monomeric FAD-containing flavoenzyme that can use NADH or NADPH as a cofactor with equal efficiency (Knox *et al.*, 1988a). The rat enzyme has a nitroreductase activity that reduces the 4-nitro group of CB1954 to 5-(aziridin-1-yl)-4-hydroxylamino-2-nitro-benzamide (Figure 1-9). This occurs by two separate two-electron reduction steps, first of the nitro group to a nitroso intermediate and then to the hydroxylamine product. Upon nitroreduction, CB1954 is converted from a weak monofunctional alkylating agent, by means of the aziridine group, to a potent difunctional alkylating species through activation of the 4-nitro group (Knox *et al.*, 1991, Knox *et al.*, 1988b, Friedlos *et al.*, 1992). The 4-hydroxylamino-2-nitro-

benzamide reduction product has been shown to form interstrand crosslinks in the DNA of Chinese hamster V79 cells co-cultured with Walker rat carcinoma cells, but cannot crosslink naked DNA alone (Knox *et al.*, 1988b). The 4-hydroxylamino product must be activated by cellular thioesters, including acetyl-coenzyme A, before it can crosslink DNA (Knox *et al.*, 1991). The proposed structure of the crosslink is shown in Figure 1-10. The 4-hydroxylamino product is, on a dose basis, 10^4- to 10^5-fold more effective at causing cell death than the prodrug alone (Roberts *et al.*, 1986). CB1954 was effective against rat tumours due to the presence of NQO1, however the equivalent human enzyme does not reduce the prodrug effectively (Chen *et al.*, 1997). The difference has been shown to be primarily due to the presence of a Gln in place of Tyr104. Mutating the human enzyme back to Tyr confers similar activity with CB1954 to the rat enzyme (Wu *et al.*, 2001). A gene therapy approach such as VDEPT would allow the insertion of genes encoding other nitroreductase enzymes including bacterial nitroreductases (Knox *et al.*, 1993b), making the prodrug therapy an option for cancer chemotherapy.

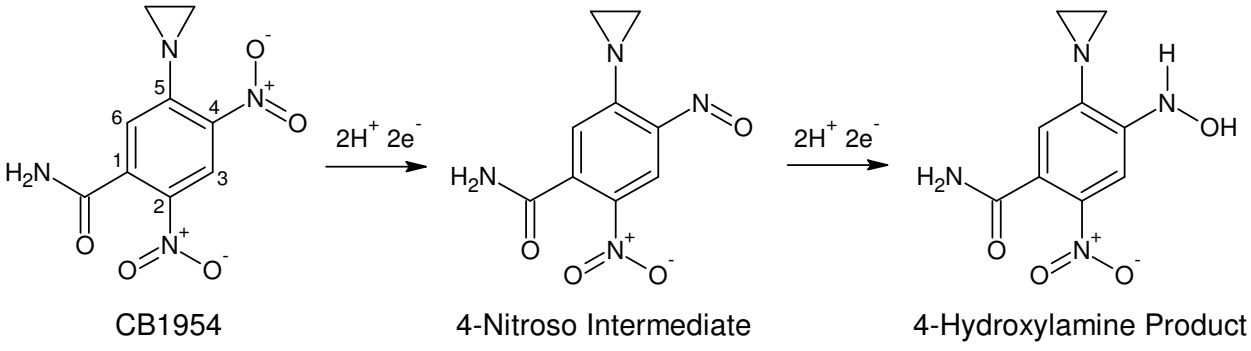


Figure 1-9: Reduction of CB1954 as catalysed by NQO1, two separate two-electron reduction steps for a total of 4 electron reduction. Figure produced using ChemSketch (ACD/labs).

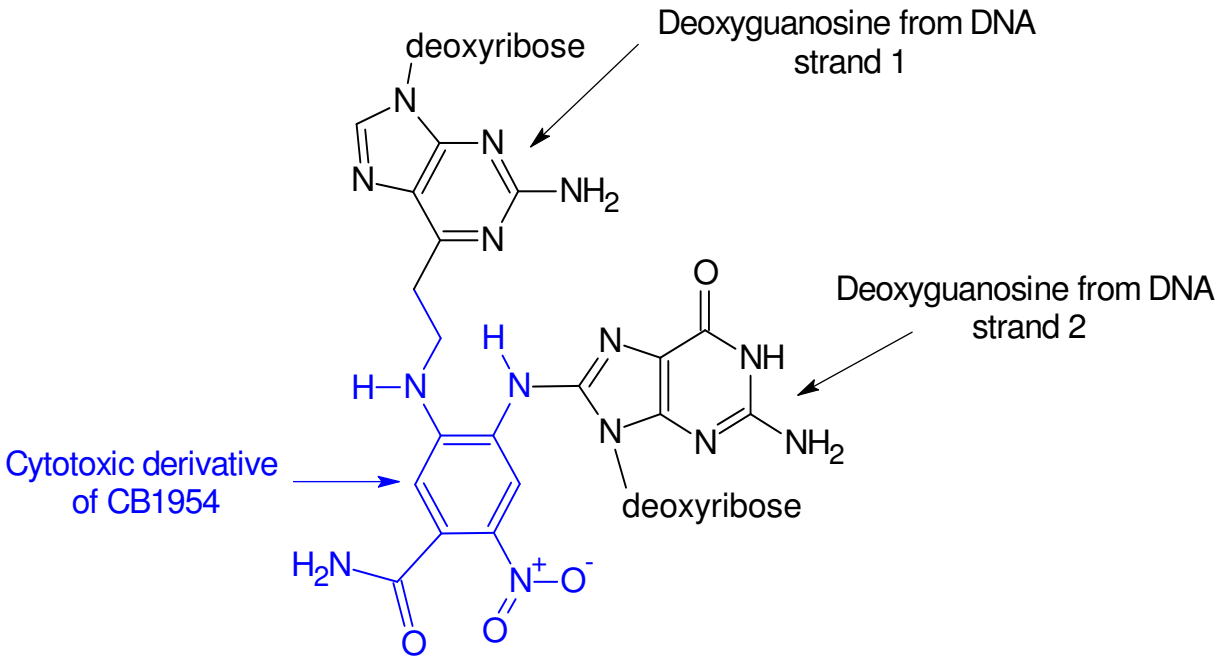


Figure 1-10: The structure of the cytotoxic crosslinking aduct formed by the reaction of the reactive 4-N-acetoxy derivative of CB1954 with DNA. Note that the two bases are not paired bases. Figure produced using ChemSketch (ACD/labs).

1.4.4.2 Nitroreductase and CB1954

It was discovered that nitroreductase-deficient *E. coli* strains showed resistance to CB1954 (Venitt and Crofton-Sleigh, 1987), and thus, that in common with NQO1, *E. coli* nitroreductases reduce CB1954 to a potent cytotoxic DNA-crosslinking agent. NfsB was isolated and found to reduce CB1954 much more rapidly than NQO1, making it a more attractive choice for gene therapy with CB1954 (Knox *et al.*, 1992).

Unlike NQO1, NfsB can reduce either (but not both) of the two nitro groups on CB1954 to the corresponding hydroxylamine derivative (Knox *et al.*, 1992) (Figure 1-11). The reduction of CB1954 by NADH and NfsB produces approximately equal proportions of the 2- and 4-hydroxylamine products (Boland *et al.*, 1991, Anlezark *et al.*, 1992), unlike DT-Diaphorase which only produced the 4-hydroxylamine (Knox *et al.*, 1988a, Knox *et al.*, 1988b). The 4-hydroxylamine product is the more cytotoxic of the two products (Knox *et al.*, 1988b, Sunters *et al.*, 1992, Helsby *et al.*, 2003), and it has been suggested that it is only this product which can form DNA crosslinks (Knox *et al.*, 1991). Both of the hydroxylamine products can be further reduced to amines *in vivo*, and the 2-amine has also shown some significant cytotoxicity (Tang *et al.*, 2005) and mediates a greater bystander effect (Helsby *et al.*, 2004) (discussed in next section). The O-acetylation of the hydroxylamine derivatives has been shown to be catalyzed by N-acetyltransferase (NAT2), and co-expression of NfsB and NAT2 increases the effectiveness of CB1954 in SKOV ovarian cancer cell lines (Mitchell and Minchin, 2008). However co-expression of NfsB and NAT2 led to a lower bystander effect and it may be that it lowers the levels of the freely diffusible hydroxylamine derivatives. Nevertheless gene therapy with both enzymes may hold a potential for future use with GDEPT.

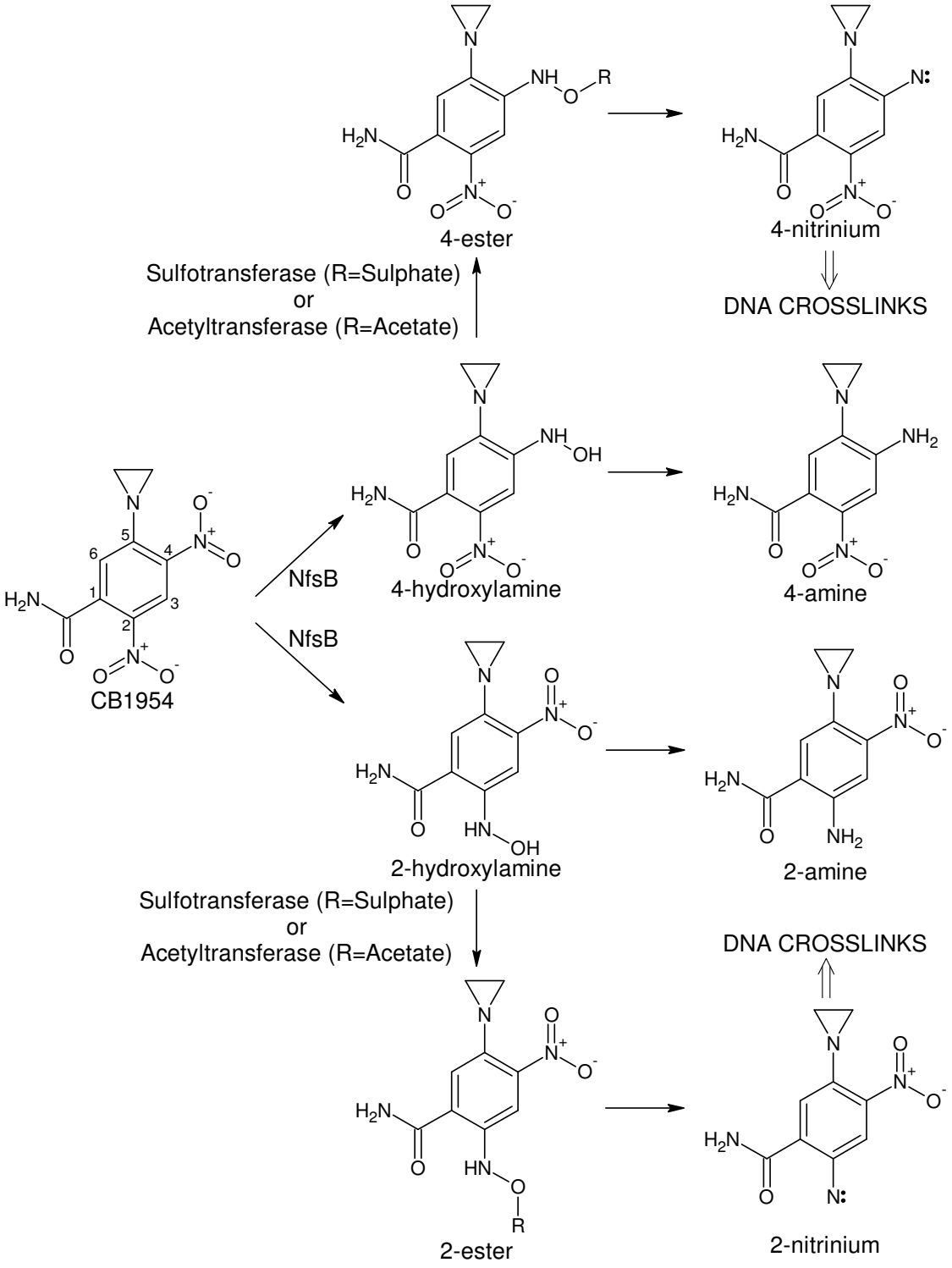


Figure 1-11: Nitroreductase-catalyzed reduction of CB1954. CB1954 can be reduced by NfsB at either nitro position (2 or 4). This produces first a nitroso (not shown) and then a hydroxylamine derivative. The hydroxylamine may be further reduced to an amine, or can go on to react with cellular thioesters to form a DNA alkylating species. Figure adapted from (Mitchell and Minchin, 2008). Figure produced using ChemSketch (ACD/labs).

1.4.4.3 Bystander effect

An essential ability of any GDEPT combination would be the ability to cause a so called 'bystander effect'. A bystander effect is the cell death of untransfected cells caused by the diffusion of cytotoxic metabolites from the transfected cells. A significant bystander effect is essential for a viable GDEPT system because only a minority of tumour cells will express the enzyme (Palmer *et al.*, 2004, Patel *et al.*, 2009). The combination of NfsB and CB1954 has shown a considerable bystander effect, both in tissue culture models (Bridgewater *et al.*, 1997, Green *et al.*, 1997, Friedlos *et al.*, 1998, McNeish *et al.*, 1998, Nishihara *et al.*, 1998, Westphal *et al.*, 2000, Wilson *et al.*, 2002, Benouchan *et al.*, 2003) and in experimental tumours (McNeish *et al.*, 1998, Djeha *et al.*, 2000, Westphal *et al.*, 2000, Wilson *et al.*, 2002, Benouchan *et al.*, 2006). The effect is due to the cell-permeable hydroxylamine derivatives of CB1954 reduction by NfsB. Of the two products it was initially thought that the 4-hydroxylamine product was responsible for the majority of the bystander effect observed, because of its high toxicity in low density cell cultures (Bridgewater *et al.*, 1997). More recently it has been shown that the 2-hydroxylamine produces a more significant bystander effect (Helsby *et al.*, 2004). Optimizing any GDEPT system purely for maximum bystander effect may not be desirable, as too large an increase in the spread of the cytotoxic agent is likely to cause an increase in unwanted systemic effects (Dachs *et al.*, 2009).

1.4.4.4 Development of the NfsB/CB1954 GDEPT combination

The enzyme prodrug combination of NfsB and CB1954 was initially suggested for its potential in ADEPT (Anlezark *et al.*, 1992, Anlezark *et al.*, 1995, Knox *et al.*, 1993b,

Knox *et al.*, 1995). Since then it has been trialled in CDEPT (Theys *et al.*, 2006), and more widely with a GDEPT/VDEPT approach. Replication-deficient retrovirus and adenovirus vectors have been used to insert the *nfsB* gene into cancer cell lines, and conferred much greater sensitivity to CB1954 in transfected cells than untransfected cells (McNeish *et al.*, 1998, Djeha *et al.*, 2000, Weedon *et al.*, 2000). *In vivo* experiments have also shown anti-tumour activity in Balb-C nude mice (Djeha *et al.*, 2000, Djeha *et al.*, 2001, Weedon *et al.*, 2000, McNeish *et al.*, 1998). There have been a number of early-stage clinical trials, first with CB1954 alone (Chung-Faye *et al.*, 2001) and second with a replication-defective adenovirus expressing NfsB (Palmer *et al.*, 2004). The trial with CB1954 alone revealed potentially dose-limiting hepatotoxicity, which may be due in part to reduction of CB1954 by native enzymes in the liver (Tang *et al.*, 2005).

The NfsB/CB1954 combination has been tested in phase I/II clinical trials in patients suffering from prostate cancer, using direct intraprostatic injection of the NfsB-expressing replication-defective adenovirus (Patel *et al.*, 2009). Some patients showed a decline in PSA (prostate-specific antigen – a tumour marker), and in a number of these this was maintained over a six month period. This suggests an anti-tumour activity. However not all patients showed a response, and the reductions seen in PSA were only partial and transient. The viral vector was well-tolerated with no dose-limiting toxicity seen at up to 1×10^{12} virus particles. This study showed that the treatment shows promise, in showing some anti-tumour activity, and warrants further study, but that the system would need further development to be effective.

1.5 Improvements to the Nitroreductase/CB1954 Combination

1.5.1 Limitations of the NfsB/CB1954 combination

The published K_m of NfsB enzyme with CB1954 is only 17.2 ± 4.8 mM (Jarrom *et al.*, 2009), whilst the peak serum concentration achievable in humans is much lower at 5-10 μM (Chung-Faye *et al.*, 2001). The clinical trial results have also suggested that improvements would be required in order to create a more effective treatment (Patel *et al.*, 2009). The combination requires greater efficacy to be viable.

1.5.2 Alternative prodrugs for nitroreductase suicide gene therapy

Although CB1954 was the first chemical proposed for use as a prodrug in nitroreductase suicide gene therapy (Anlezark *et al.*, 1992), many other prodrugs have been studied for their potential for use alongside nitroreductase. Four classes of prodrugs for use with nitroreductase have been discussed elsewhere in detail; dinitrobenzamide mustards, dinitroaziridinylbenzamides (CB1954 analogues), 4-nitrobenzylcarbamates and nitroindolines (Denny, 2002). Other classes of potential prodrug include acyclic nitrobenzylphosphoramide mustards (Jiang *et al.*, 2006), other nitroaromatics and quinones (Bailey *et al.*, 1996).

1.5.2.1 Dinitrobenzamide Mustards

Aromatic nitrogen mustard alkylating agents, such as melphalan and chlorambucil, have been used in isolation as anti-cancer drugs in chemotherapy (Pratt and Ruddon,

1979). They kill cells by alkylating bases of DNA, causing inter- and intra-strand crosslinks and even strand breaks. Although they have a greater effect on proliferating cells (crucially including rapidly-dividing tumour cells), the dosages that can be used clinically are still limited due to their effect on non-cancerous cells. Aromatic nitrogen mustards can have their activity modified via altering ring substituents. They therefore have a potential as prodrugs in cancer gene therapy, as their substituents could be enzymatically altered from electron-withdrawing to electron-donating, activating the cytotoxic properties of the prodrug (Palmer *et al.*, 1990). SN23862 is the bischloroethylamino mustard analogue of CB1954, and can be reduced by NfsB (Michael *et al.*, 1994) (Figure 1-12).

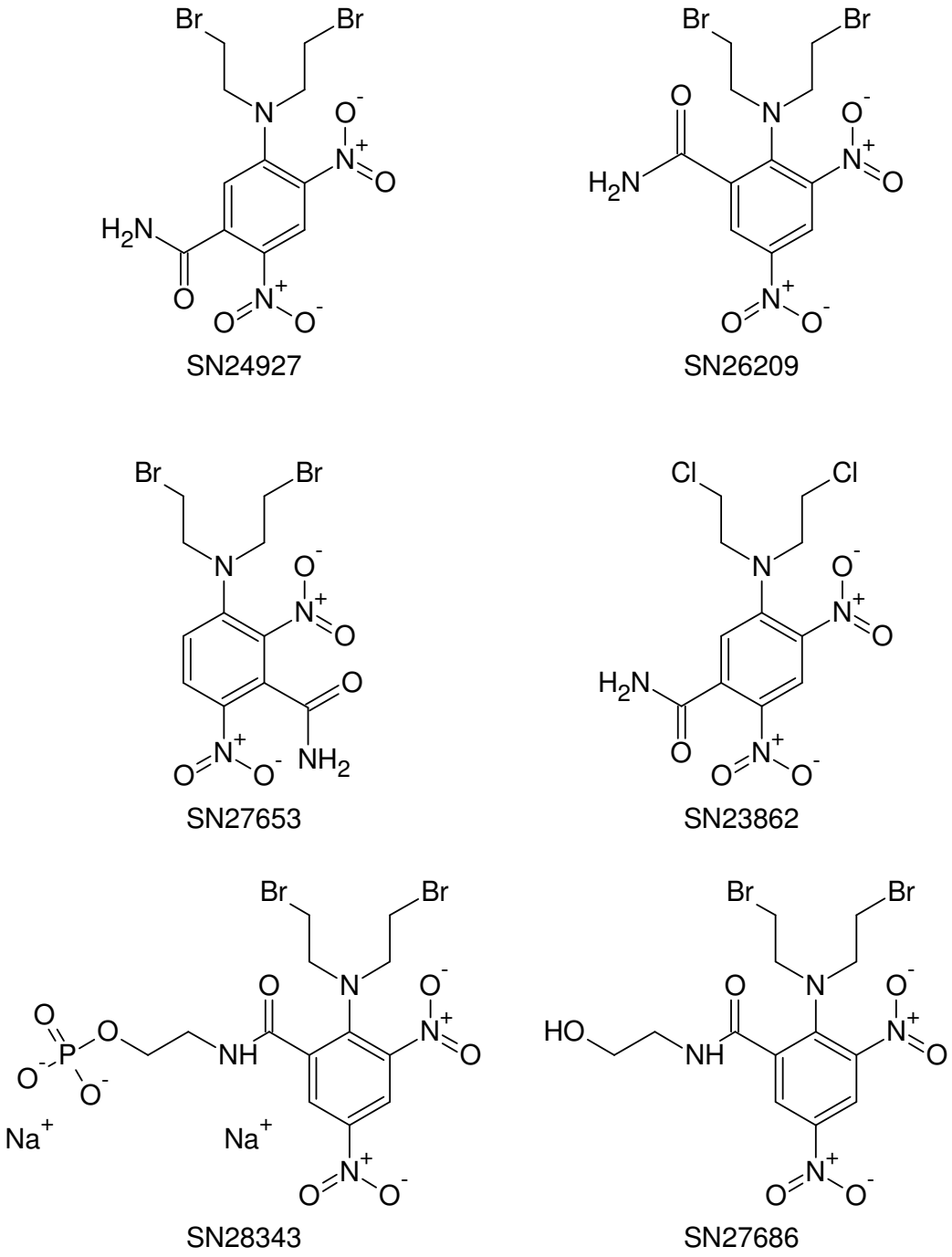


Figure 1-12: Structures of the most promising alternative dinitrobenzamide prodrugs to date. SN23862 is the bis-chloroethyl analogue of CB1954 and has been most extensively studied as an alternative (Michael *et al.*, 1994). SN27686 is the most active of a new series of related prodrugs, and its phosphate ester, SN28343, is its pre-prodrug with greater aqueous solubility (Singleton *et al.*, 2007). The phosphate is removed by phosphatises in the serum. The others have also been tested with NfsA. Figure produced using ChemSketch (ACD/labs).

Other dinitrobenzamide mustards have been touted as potential alternative prodrugs with nitroreductase (Anlezark *et al.*, 1995, Atwell *et al.*, 1996, Atwell *et al.*, 2007, Friedlos *et al.*, 1997). The first of the dinitrobenzamide mustards to be developed for PAE were based upon adding or changing the substituents of CB1954. Another advantage of SN23862 and related prodrugs is that they are not activated by mammalian DT Diaphorases (Anlezark *et al.*, 1995, Palmer *et al.*, 1992, Palmer *et al.*, 1994), leading to potentially fewer side effects. Unlike CB1954, SN23862 is only activated by NfsB at its 2-nitro position (Anlezark *et al.*, 1995) which, it has been suggested from preliminary modelling, is because the larger mustard group restricts entry of the drug into NfsB's active site (Parkinson *et al.*, 2000). However reduction of the 4-nitro group is fully activating as SN23862 is already a difunctional alkylating agent (through its mustard moiety) and the activation by NTR releases electrons activating the DNA-alkylating ability of the drug. Related prodrugs have been created by replacing the 2-nitro group with non-activating groups of various electrostatic properties (Atwell *et al.*, 1996), substituted amides, differently-located nitro groups and different halogen mustards (Friedlos *et al.*, 1997). To date the most promising of these dinitrobenzamide mustards, have longer carbon chains on the amide moiety and make the prodrug more lipophilic and thus more able to pass across membranes (Singleton *et al.*, 2007). The water soluble pre-prodrug SN28343 is transformed *in vivo* by serum phosphatases into the prodrug SN27686, which has shown improved tissue penetration and bystander effect upon activation by NfsB relative to CB1954 (both shown in Figure 1-12).

1.5.2.2 4-Nitrobenzyl Carbamates

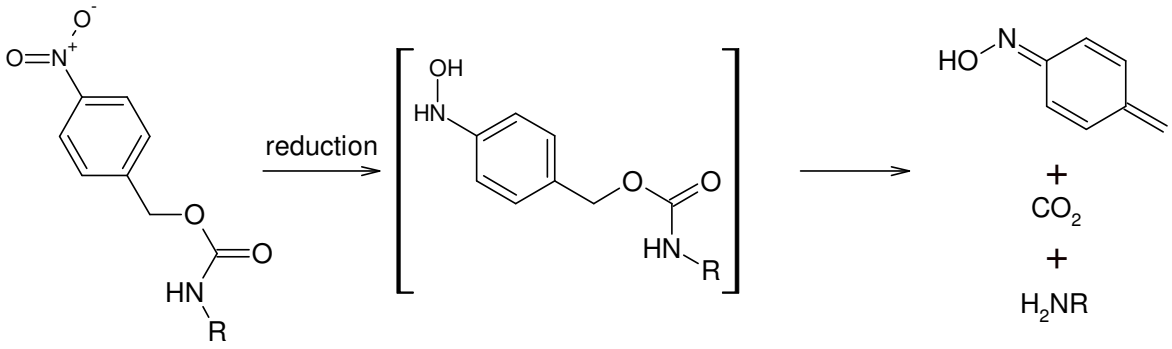


Figure 1-13: Fragmentation reaction of 4-nitrobenzyl carbamates by nitroreductase (Denny, 2002). The H_2NR is the activated drug. Figure produced using ChemSketch (ACD/labs).

4-Nitrobenzyl carbamate prodrugs have been shown to be effective substrates for NfsB (Mauger *et al.*, 1994). Chemotherapeutic drugs can be coupled to a 4-Nitrobenzyl carbamate, and reduction by nitroreductase causes a fragmentation (See Figure 1-13) leaving the active drug (Hay *et al.*, 1999a). Carbamate prodrugs have been synthesised and tested for a variety of classes of chemotherapeutic drugs, including aniline mustards (Mauger *et al.*, 1994), enediynes (Hay *et al.*, 1995, Hay *et al.*, 1999b), pyrrolobenzodiazepines (Sagnou *et al.*, 2000) and duocarmycin analogues (Hay *et al.*, 1999a). Drugs trialled in this system include actinomycin D, mitomycin D and doxorubicin (Mauger *et al.*, 1994). Most studies have focussed on ADEPT rather than GDEPT/VDEPT but, except with hydrophilic drugs, they should work in a VDEPT approach.

1.5.2.3 Nitrobenzylphosphoramide Mustards

Another related class of prodrugs are the acyclic nitrobenzylphosphoramide mustards. The origins of this class were nitrobenzene analogues fused to cyclophosphamide (Li *et al.*, 2003). A series of analogues were generated with the nitro group *para* to the benzylic carbon, the most promising of which, LH7 (Jiang *et al.*, 2006) (Figure 1-14), has been studied more extensively with NfsB and its most promising mutants (Jaberipour *et al.*, 2010). Subsequent studies examined the effect of the position of the nitro group on the benzene ring, and the effect of subsequent substitutions (Hu *et al.*, 2011). It was found that halogens positioned ortho to the benzyl carbon showed the greatest improvement in activity in V79 and SKOV3 cancer cell lines, however discrepancies were seen when tested in bacterial cell lines. It has been proposed that this group of prodrugs are DNA cross-linking agents, and involve cleavage of the phosphoramide mustard (Jiang *et al.*, 2006).

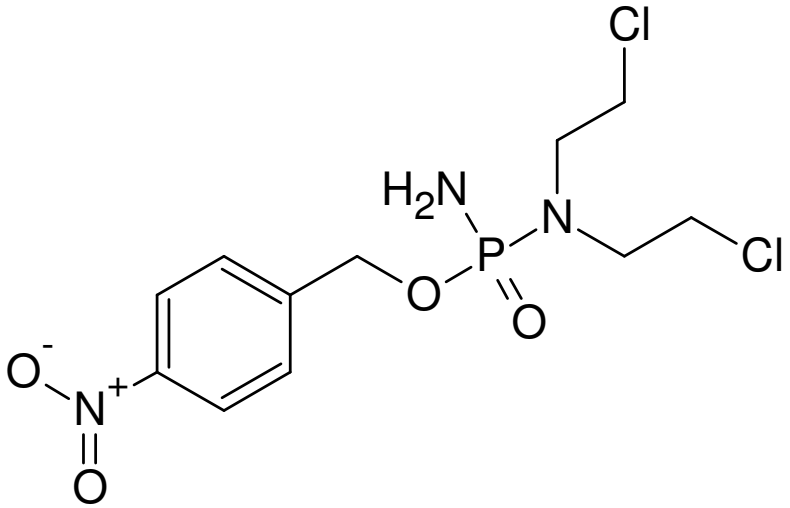


Figure 1-14: The structure of the nitrobenzylphosphoramide mustard LH7 (Jiang *et al.*, 2006). Figure produced using ChemSketch (ACD/labs).

1.5.2.4 Other Classes of Alternative Prodrugs

A few other classes of drug have been tested for use with *E. coli* NfsB. These include nitroindolines, which have shown some promise in tests (Tercel *et al.*, 1996). These work by binding to the minor grove of DNA, but are not greatly activated by nitroreduction. Of other nitroaromatic compounds that have been investigated for potential use as a prodrug, only the antibiotic nitrofurazone has shown any promise at all (Anlezark *et al.*, 1995, Bailey *et al.*, 1996). Quinone substrates have so far proven ineffective (Bailey *et al.*, 1996).

1.5.3 Improving the enzyme for the VDEPT combination

1.5.3.1 Mutating NfsB

An alternative to improving the prodrug is to change the enzyme used for the combination. Mutants of NfsB have been developed which confer greater sensitivity to CB1954 in both human and bacterial cell lines (Grove *et al.*, 2003, Guise *et al.*, 2007, Jaberipour *et al.*, 2010). These will be discussed in more detail later. The gene has also been mutated to optimise its codon usage in mammalian systems, since this differs between prokaryotes and eukaryotes (Grohmann *et al.*, 2009).

1.5.3.2 NfsA as a Prodrug Activating Enzyme

Alternative enzymes could be used in combination with CB1954; indeed the major nitroreductase of *E. coli*, NfsA, has been suggested as an alternative enzyme for VDEPT with CB1954 as it confers between a 3.5 and 8-fold increase in sensitivity to CB1954 in human cell lines compared to NfsB (Vass *et al.*, 2009). NfsA was also

found to be the most active enzyme for the prodrug in a study comparing multiple potential nitroreductase enzymes from *E. coli* (Prosser *et al.*, 2010a).

1.5.3.3 Other Alternative Enzymes

Nitroreductase enzymes from other organisms have also been tested in combination with CB1954. The mammalian enzyme NRH:quinone oxidoreductase 2 (NQO2), which shares 49% sequence identity with NQO1 (DT-Diaphorase) although it is 43 amino acids shorter (both enzymes share less than 5% identity with either NfsA or NfsB, and bind FAD, with the isoalloxazine ring in an opposite conformation), has been suggested as a potential prodrug activating enzyme with CB1954 (Knox *et al.*, 2000, Knox *et al.*, 2003). It is of particular interest as the enzyme is over-expressed in several cancer cell lines (Knox *et al.*, 2000). NQO2 reduces CB1954 3000 times faster than NQO1, and at a similar rate to NfsB with a lower K_m (263 μM for NQO2 and 900 μM for NfsB (Race *et al.*, 2007)). It also reduces CB1954 only at the 4-nitro position (more cytotoxic but reduced bystander effect), unlike NfsB which reduces at both 4-nitro and 2-nitro. The structure of NQO2 bound to CB1954 has been determined by X-ray crystallography, but it was not possible to distinguish between two possible orientations of the prodrug (AbuKhader *et al.*, 2007). Subsequent NMR studies have determined the orientation of the ligand conclusively (Ludwig *et al.*, 2008) (

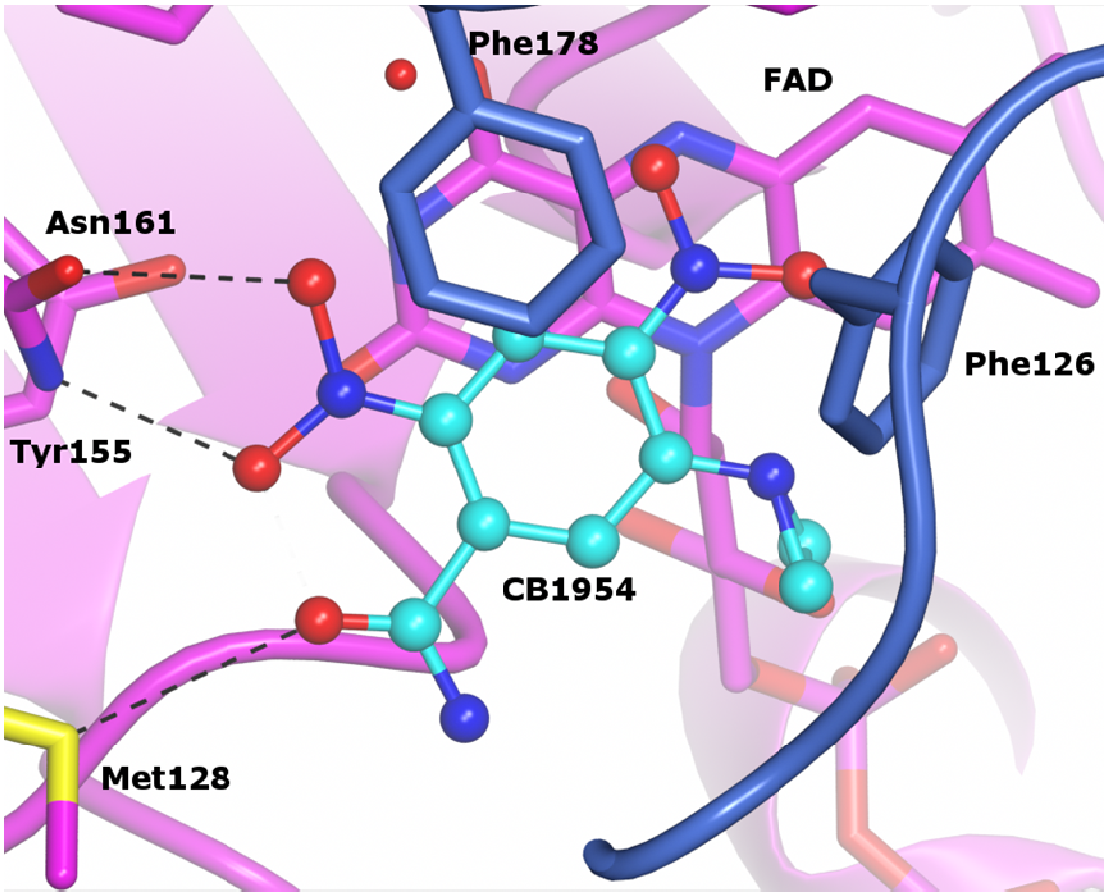


Figure 1-15). A phase I clinical trial has been undertaken with CB1954 and the NQO2 co-substrate EP0152R, showing that the treatment is not harmful (Middleton *et al.*, 2010). NQO2 requires co-administration of a synthetic cofactor, (EP015R in the study mentioned here), unlike NfsB that utilises NAD(P)H, which is found in the cell.

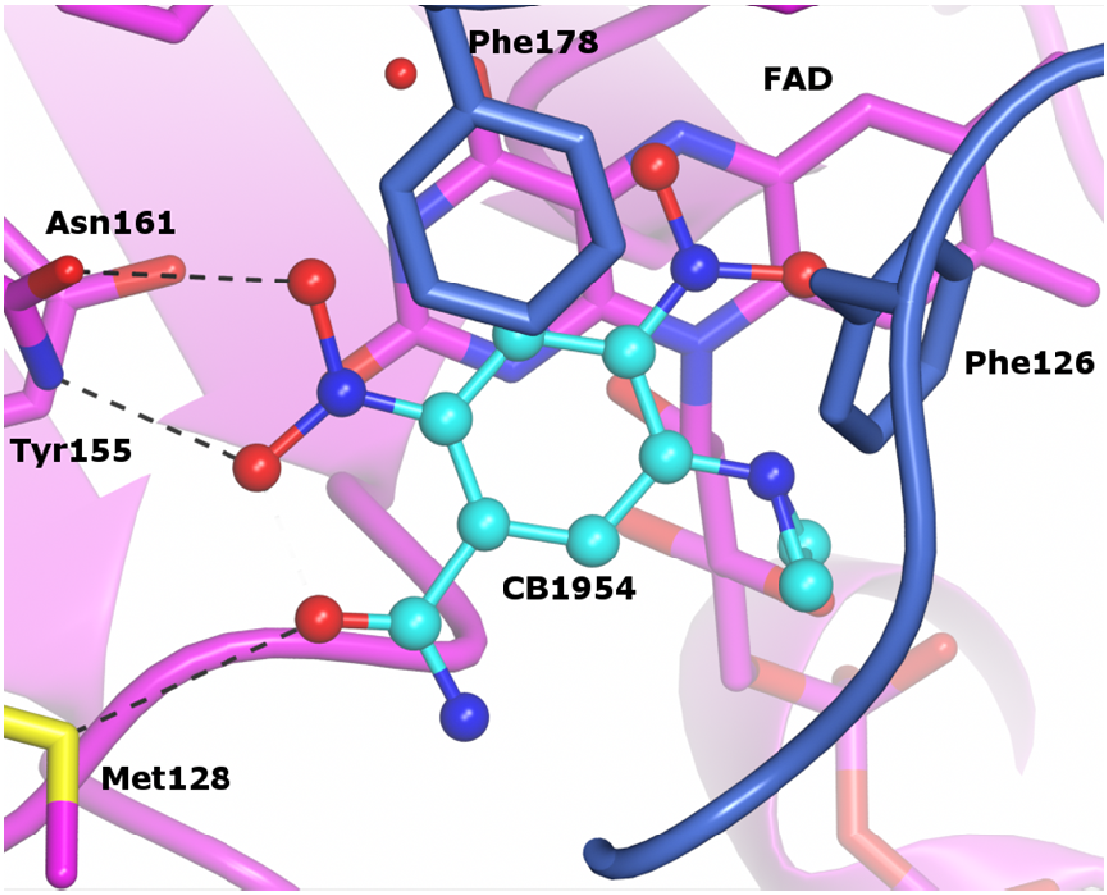


Figure 1-15: CB1954 (light blue –Figure 1-9) bound in the active site of NQO2 (pdb I.D. 1ZX1 (Ludwig *et al.*, 2008)). The black dashed lines represent H-bonds. Figure produced using CCP4 Molecular Graphics.

A summary of some of the most promising alternative enzymes is shown in Table 1-2. Several studies have compared various enzymes for use with CB1954 (Theys *et al.*, 2006, Prosser *et al.*, 2010a, Prosser *et al.*, 2010b). Potential nitroreductase-like enzymes from *E. coli* have been tested for activity with CB1954 *in vivo* by exploiting the bacterial SOS response upon DNA damage, though monitoring expression of β-galactosidase activity of an SOS response promoter (Prosser *et al.*, 2010a). This study suggested only NfsA had a greater potential then NfsB.

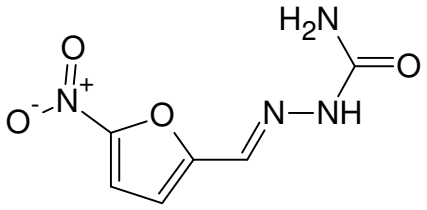
Enzyme ***Organism***	**% Similarity (*identity*) to NfsB**	**% Similarity (*identity*) to NfsA**	**Improvement Seen relative to WT NfsB**	**Reference**
NQO2 ***Human***	3.8 (*2.2*)	5.4 (*3.5*)	Similar rate, with lower K_m (263 µM versus 900 µM)	(Knox *et al.*, 2003)
YfkO ***Bacillus licheniformis***	43.1 (*23.4*)	28.6 (*15.5*)	Improvements claimed but no relevant data is shown	(Emptage *et al.*, 2009)
YfkO ***Bacillus subtillis***	43.6 (*28.4*)	30.7 (*18.6*)	Less effective than NfsB	(Prosser *et al.*, 2010b)
NfsA ***Klebsiella pneumoniea***	23.6 (*12.6*)	92.1 (*83.8*)	Less effective than NfsB	(Prosser *et al.*, 2010b)
NbzA ***Pseudomonas pseudoalcaligenes***	33.5 (*19.8*)	27.1 (*18.0*)	K_m of 11.7 µM is claimed but no kinetic data is shown	(Berne *et al.*, 2006)
NemA (N-ethylmaleimide reductase) ***E. coli***	16.8 (*9.5*)	4.3 (*3.1*)	Less effective than NfsB	(Prosser *et al.*, 2010a)
AzoR (azoreductase) ***E. coli***	24.9 (15.3)	4.6 (*2.7*)	Less effective than NfsB	(Prosser *et al.*, 2010a)
YieF (chromate reductase) ***E. coli***	13.7 (*6.4*)	25.6 (*14.8*)	Some initial promise shown but more study would be required	(Barak *et al.*, 2006)

Table 1-2: Summary of some of the most promising alternative enzymes for use with CB1954 in GDEPT.

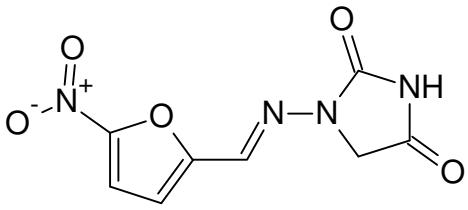
1.6 Properties and Structures of the *E. coli* Nitroreductase Enzymes

1.6.1 Mechanism of Reduction by *E. coli* Nitroreductases

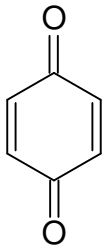
Both NfsA and NfsB catalyse the flavin-mediated reduction of a broad range of nitroaromatic and quinone derivatives (Zenno *et al.*, 1996c, Zenno *et al.*, 1996a); for some examples see Figure 1-16. Studies have suggested that both enzymes follow a substituted enzyme or 'ping-pong' mechanism (Anlezark *et al.*, 1992, Zenno *et al.*, 1996a, Zenno *et al.*, 1996c). A ping-pong mechanism is where an enzyme, catalysing a bi-substrate reaction, interacts with only one substrate at a time, and switches between a modified and unmodified form (detailed in the case of NfsB and CB1954 in Figure 1-17). The mechanism has two distinct steps; the first step, called the 'ping' step, is where the first substrate interacts with the native enzyme to create a modified form of the enzyme. In second, or 'pong', step the modified enzyme interacts with the second substrate, forming the products of the reaction and reforming the native form of the enzyme.



Nitrofurazone

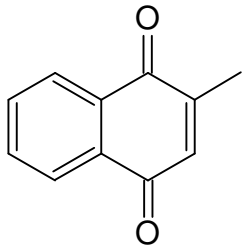


Nitrofurantoin

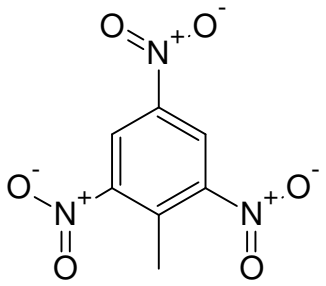


Quinone

(1-4, Benzoquinone)



Menadione



TNT (2,4,6-Trinitrotoluene)

Figure 1-16: Chemical structures of some examples of the known substrates of NfsA and NfsB (Zenno *et al.*, 1996a, Zenno *et al.*, 1996c). Figure produced using ChemSketch (ACD/labs).

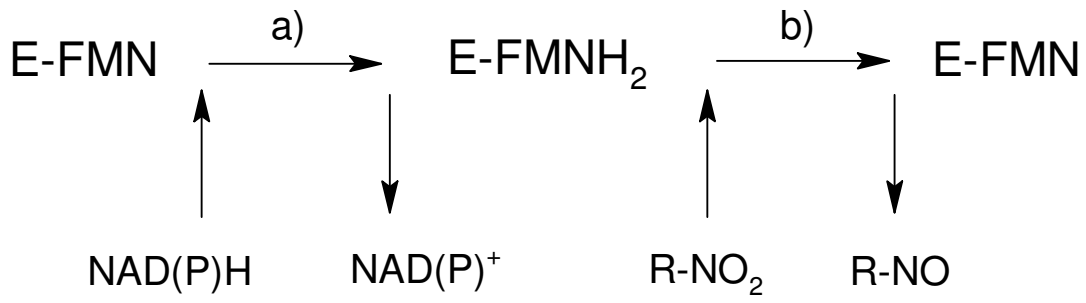


Figure 1-17: Summary of substituted ping-pong bi-bi reaction catalysed by the FMN bound nitroreductase enzymes.

With NfsA and NfsB, the enzyme is a complex with FMN, and the modification is of the FMN rather than the enzyme, in common with most flavoenzymes (E-FMN in Figure 1-17). The native form of the enzyme contains the FMN in its oxidised form. The 'ping' step (labelled a) in Figure 1-17) for the enzymes is where the enzyme-FMN complex interacts with either NADH or NADPH, and the incoming cofactor donates two electrons, reducing the FMN. This produces the modified intermediate form of the enzyme-FMN complex, with reduced FMN (E-$FMNH_2$), which goes on to interact with the nitroaromatic or quinone substrate in the 'pong' part of the reaction. In this second interaction (labelled b) in Figure 1-17), the electrons are transferred from the $FMNH_2$ to the second substrate, reducing it and reforming the native form of the enzyme bound to oxidised FMN (Anlezark *et al.*, 1992).

The initial assessment of a ping-pong mechanism for NfsB was made based upon kinetic observations (Anlezark *et al.*, 1992), which were backed up by further more in depth kinetic and inhibition studies (Zenno *et al.*, 1996c, Race *et al.*, 2005). Similar

kinetic studies on an *E. cloacae* nitroreductase that is highly homologous to NfsB also supported a ping-pong mechanism (Koder and Miller, 1998). Crystallographic studies of NfsA (Kobori *et al.*, 2001), NfsB (Lovering *et al.*, 2001) and the *E. cloacae* nitroreductase (Haynes *et al.*, 2002) support this mechanism, as the active sites are not able to accommodate both substrates, NAD(P)H and the nitroaromatic/quinone, simultaneously. Crystal structures have been determined of the reduced form of FMN complexed with NfsB (Parkinson *et al.*, 2000, Johansson *et al.*, 2003) and with the *E. cloacae* nitroreductase (Haynes *et al.*, 2002), further supporting the proposed mechanism. A comparison of steady-state and stopped-flow kinetic rates suggest that the rate-determining step is the release of the product for the reduction of CB1954 and NFZ by WT-NfsB, as the maximum rate observed in the steady-state is lower than that of either step in the stopped-flow (Jarrom *et al.*, 2009).

The reduction of quinones by the nitroreductases occurs via a single two-electron reduction step to an alcohol group, while 4 electrons are needed to reduce a nitro group of nitroaromatics. The reduction of nitro groups occurs by two successive two-electron steps, as shown for CB1954 in Figure 1-11 and as discussed in Section 1.4.4.2. The hydroxylamine is readily reduced *in vivo* to an amine, but this is not seen *in vitro*. The reduction of the nitro to the nitroso has been shown to be the rate-limiting step for the reduction catalyzed by NfsB (Race *et al.*, 2005). The second step is quick even without the enzyme and whilst it has been shown that it can occur faster in the presence of the enzyme, it may also just involve reduction by free NAD(P)H in solution (Race *et al.*, 2005).

1.6.1.1 Hydride or Electron Transfer?

For two-electron reduction of CB1954 catalyzed by NfsA and NfsB a total of two electrons and two protons are transferred upon reduction of the nitro to the nitroso. Both electrons must come from the FMN and one proton must come from solution, but the other proton could come from either FMN or solution (Christofferson and Wilkie, 2009). The reaction could thus occur by two mechanisms: hydride transfer or electron transfer. In hydride transfer a hydride (two electrons and one proton) is transferred directly from the FMN to the nitro group of the substrate, either before (Figure 1-18) or after (Figure 1-19) a proton from solution. For hydride transfer either the nitrogen or either oxygen from the nitro group must come within van der Waals contact of the HN5 of the FMN (for numbering see Figure 1-4 c)). As both CB1954 and FMN have delocalised aromatic rings, an electron transfer mechanism (where only the electrons are transferred from the FMN and the proton comes from solution Figure 1-20) only requires some part of the CB1954 wave function to be within range to interact with the wave function of the FMN flavin ring system.

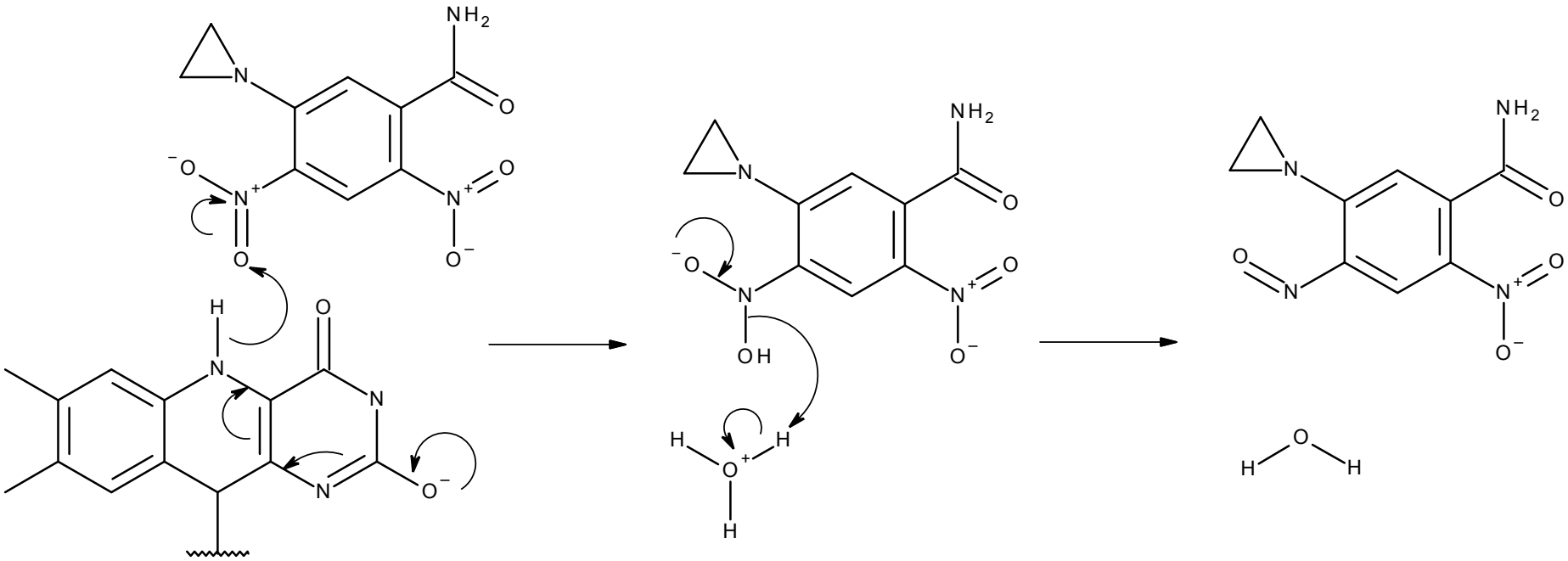


Figure 1-18: Net hydride transfer followed by proton transfer from solution. A possible mechanism for the reduction of CB1954 by reduced FMN nitroreductase complex. Figure produced using Accelrys® Draw, Accelrys Software Inc.

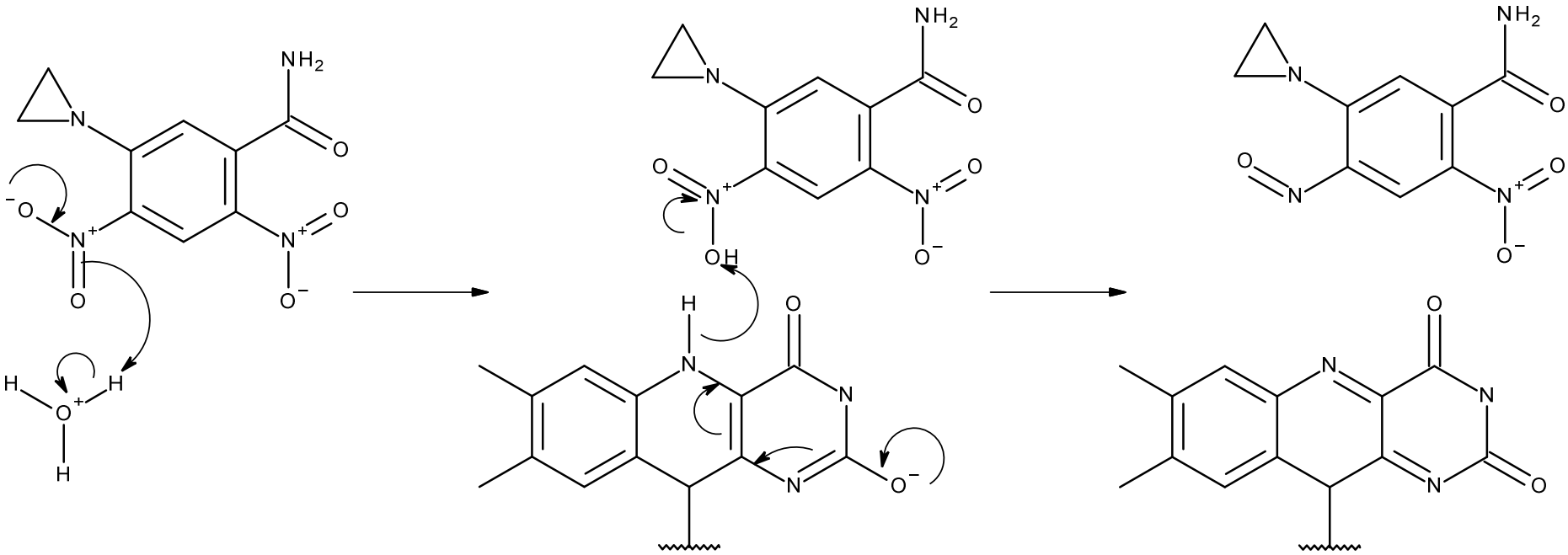


Figure 1-19: Proton transfer from solution followed by net hydride transfer. A possible mechanism for the reduction of CB1954 by reduced FMN nitroreductase complex. Figure produced using Accelrys® Draw, Accelrys Software Inc.

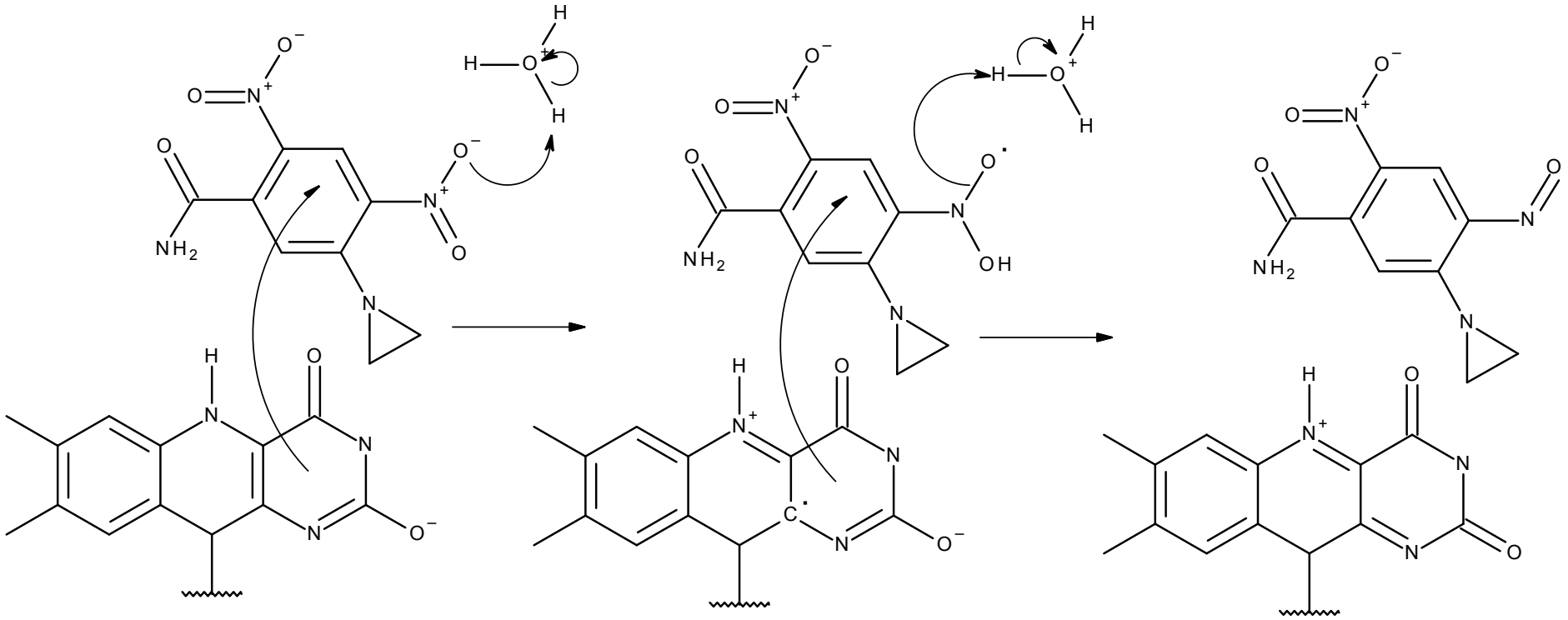


Figure 1-20: Electron transfer followed by proton transfer in two distinct steps. A possible mechanism for the reduction of CB1954 by reduced FMN nitroreductase complex. Figure produced using Accelrys® Draw, Accelrys Software Inc.

For transfer of a hydride the reactive centres must be within van der Waals contact (~3.6 Å), however electron transfer only requires some part of the wave functions of the the electron donor and electron acceptor to overlap allowing it to occur over much greater distances (Moser et al., 2010). The electrons can 'quantum tunnel' through an energy barrier that would classically prevent the transfer, in this case allowing transfer over larger distances. Edge to edge distances (of the two conjugated systems) of less than 14 Å would allow the transfers to occur at a rate that is not rate-limiting, as the rate would decrease exponentially with increasing distance (Page et al., 2003). Thus unlike hydride transfer, electron transfer would not require a specific productive binding orientation and could occur over longer distances in the protein environment.

Calculations of reactions of CB1954 and $FMNH_2$ in the gas phase have shown that either net hydride transfer or electron transfer are thermodynamically feasible reactions with CB1954, but molecular dynamics calculations of the protein bound to CB1954 in solution suggests electron transfer with CB1954 is the more likely (Christofferson and Wilkie, 2009). The crystal structure of NfsB has been determined with nicotinic acid, an analogue of the nicotinamide group of NAD(P)H, bound in the active site (Lovering *et al.*, 2001) (Figure 1-21). The nicotinate is bound with what would be the nicotinamide hydride donor C4 atom directly over the FMN N5, with the ideal geometry and distance for hydride transfer, implying that the reduction of the flavin by NAD(P)H occurs by hydride transfer, which would be expected for NAD(P)H (Fraaije and Mattevi, 2000). It is likely that quinone reduction also occurs by hydride transfer. Another crystal structure of NfsB contains nitrofurazone bound in the active site, however its nitro group faces out into solution and its amide group over

the flavin ring system (Race *et al.*, 2005) (Figure 1-22). If the mechanism of reduction were hydride transfer then this structure would be unproductive as the nitro group is not within range of the N5. However this structure is compatible with the electron transfer mechanism as any part of the conjugated system could be reduced. These structures are of the oxidised enzyme, and so are not of active complexes, and there may be different binding preferences for the oxidised and reduced forms of the enzyme (Race *et al.*, 2005, Barna *et al.*, 2001, Haynes *et al.*, 2002). To conclude it is still not possible to be certain which mechanism the reduction of CB1954 occurs by.

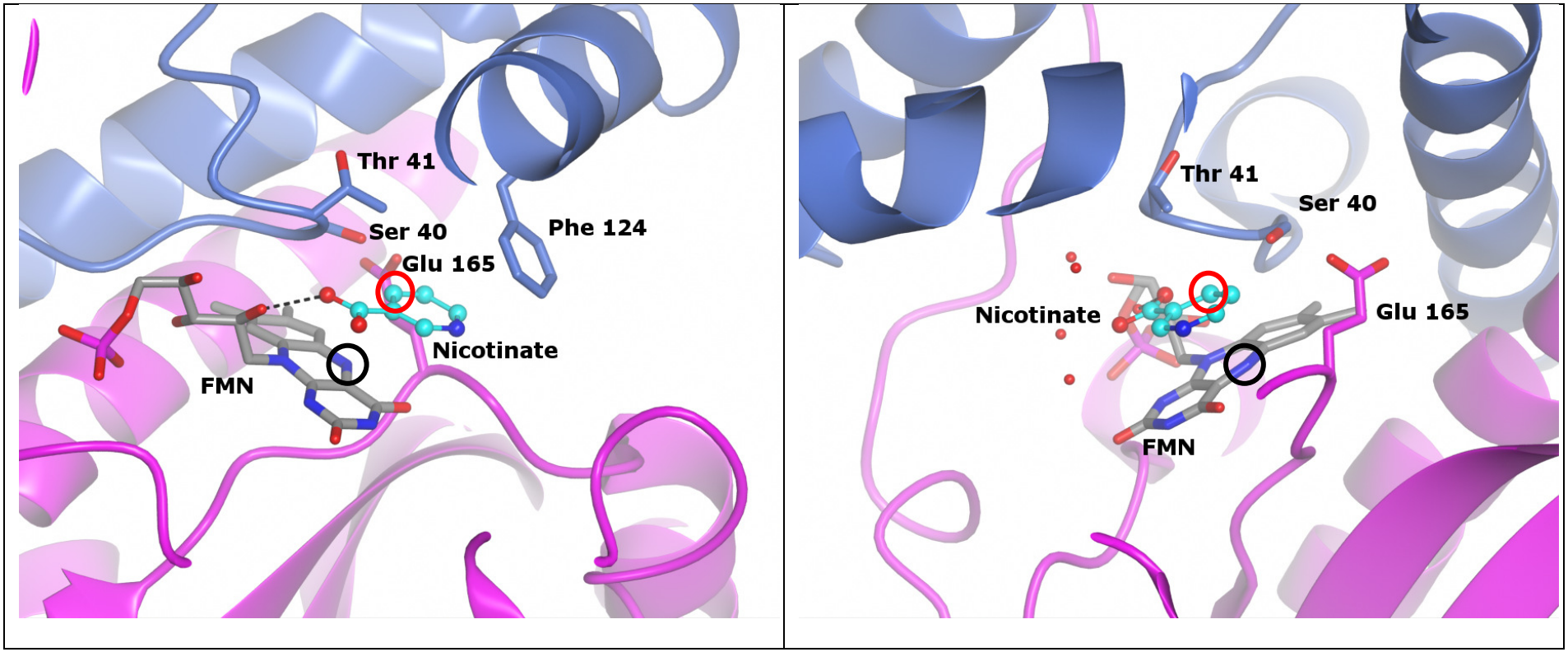


Figure 1-21: Nicotinate bound in the active site of WT-NfsB (pdb I.D. 1ICR (Lovering *et al.*, 2001)). The two images are 90° apart. The reactive N5 nitrogen of the FMN is circled in black and the carbon equivalent to the atom in NAD(P)H that donates the proton is circled in red. Figures produced using CCP4 Molecular Graphics.

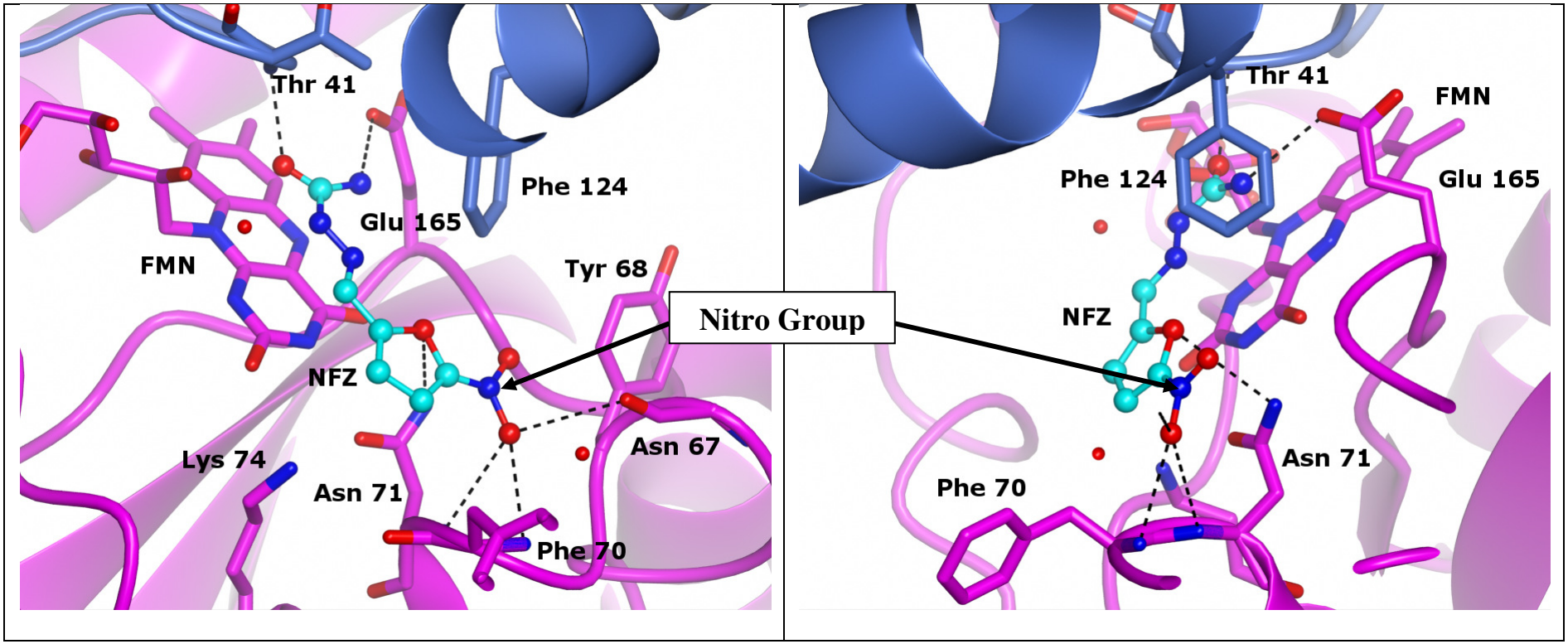


Figure 1-22: Nitrofurazone (light blue) bound in the active site of WT-NfsB (pdb I.D. 1YKI (Race *et al.*, 2005). Two images are 90° apart. H-bonds are shown as black dashed lines. The NFZ nitro group (indicated) is bound away from the catalytic centre of the FMN ring system. Figures produced using CCP4 Molecular Graphics.

1.6.2 Properties of NfsB

NfsB has been extensively studied, focussing on its potential as a prodrug activating enzyme (PAE) in cancer gene therapy. NfsB is a homodimer with each monomer containing 217 amino acid residues. It was first isolated and purified by Anlezark *et al.* in 1992, where it was revealed to be an FMN-containing flavoprotein of 24 kDa (Anlezark *et al.*, 1992). NfsB can utilise either NADH or NADPH as a cofactor. Analysis of the amino acid sequence revealed that NfsB has extensive sequence homology with both the "classical nitroreductase" of *Salmonella typhimurium* (Watanabe *et al.*, 1990) and an inducible nitroreductase found in *Enterobacter cloacae* (Bryant *et al.*, 1991). NfsB shares 88.5% identity in terms of nucleotide sequence with the two enzymes (Michael *et al.*, 1994).

1.6.2.1 Kinetics of NfsB

The kinetics of NfsB catalysis has been extensively categorised with CB1954 and other substrates, as well as some inhibitors. Due to the low solubility of many of the prodrugs it is not possible to get a full Michaelis-Menten curve, and so the initial slope of the curve, k_{cat}/K_m (the specificity constant), is used to compare the fit of substrate for enzyme. NfsB was initially reported to have a K_m of 862±145 μM (Anlezark *et al.*, 1992) making it a poor binding substrate. A later study, estimating the global kinetics (simultaneous fitting on multiple parameters, in this case the rate and concentrations of both substrates), found the k_{cat} and K_m to be ~20 fold higher (Jarrom *et al.*, 2009), but found the k_{cat}/K_m ratio to be similar to previous studies (~0.007 $\mu M^{-1} s^{-1}$) (Anlezark *et al.*, 1992). K_m's have also been determined for nitrofurazone (K_m = 1850 ± 400 μM) and menadione (K_m = 150 ± 15 μM) (Race *et al.*, 2005, Jarrom *et al.*, 2009). NfsB

shows a small 2-fold preference for NADH over NADPH for a cofactor in terms of k_{cat}/K_m (Vass *et al.*, 2009).

1.6.2.2 Structure of NfsB

1.6.2.2.1 Overall Structure of NfsB

The structure of NfsB has been determined by several groups using X-ray crystallography (Parkinson *et al.*, 2000, Lovering *et al.*, 2001, Skelly *et al.*, 2001, Johansson *et al.*, 2003), and a number of crystal forms have been described. All of the structures of NfsB currently available in the protein data bank have a common tertiary and secondary structure. The overall structure of the NfsB dimer is shown in Figure 1-23. Each monomer contains 5 β-sheets and 10 α-helices, with a central core of anti-parallel β-sheets surrounded by the α-helices. There is an extensive dimer interface, with 42 hydrogen-bonds and 418 non-bonded contacts, involving approximately 60 residues of each monomer (PDEsum - (Laskowski, 2009)). The two FMN-containing active sites are found at the dimer interface and residues from both monomers interact with substrate (Lovering *et al.*, 2001). The two active sites are identical in terms of the protein side chain positions.

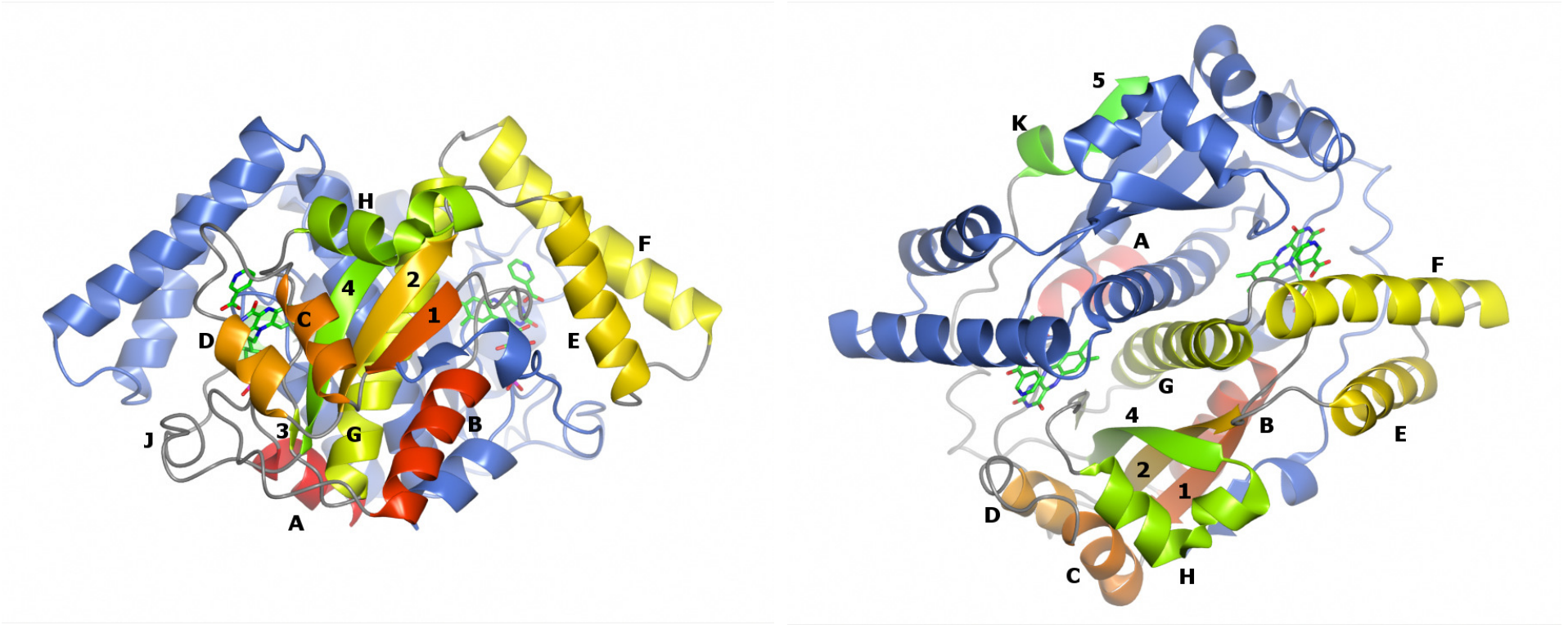


Figure 1-23: The Structure of NfsB from pdb I.D. 1ICR (Lovering *et al.*, 2001). The FMN prosthetic group and nicotinate ligand are shown as green sticks and give an indication of the position of the active site. The secondary structure elements are labelled for one half of the dimer, whilst the other half is shown in blue. The α-helices are labelled A - K and the β-sheets are labelled 1 -5. Images separated by 90°. Figure produced using CCP4 Molecular Graphics.

1.6.2.2.2 FMN Binding Site of NfsB

The FMN is tightly bound to the protein in both active sites; the binding site is shown in Figure 1-24. There are three areas of interaction with the protein, two of H-bonding with the FMN ribityl and phosphate, and the carbonyl containing ring of the isoalloxazine ring system. And finally the flavin ring system is bound primarily through hydrophobic interactions. In total there are 10 H-bonds from the flavin to the protein, with nine to one monomer and one to the other (the backbone nitrogen of Ser 39 to the ribityl O3' alcohol). The phosphate is bound in a phosphate binding pocket formed mostly of positively charged residues, comprising of Arg 10, Lys 205 and Arg 207, with additional H-bonds from the backbone and sidechain of Ser 12. The side chains of Lys 14, Lys 74 and Asn 71 all interact with the carbonyls and nitrogen of the final flavin ring as does the backbone of Gly 166. All the flavin rings some some degree of butterflying. The crystals were all formed of oxidised flavin (hence the yellow crystals) but the butterflying could be due to photoreduction in the X-ray beam. The crystal structure of the highly related nitroreductase from *Enterobacter cloacae*, has been solved in both oxidised and reduced states (Haynes et al., 2002). They see a change in the butterfly angle from 16° to 24° upon 2-electron reduction. The 16° is similar to the angle seen with all NfsB structures, and this slightly bent conformation would favour reduction of the flavin.

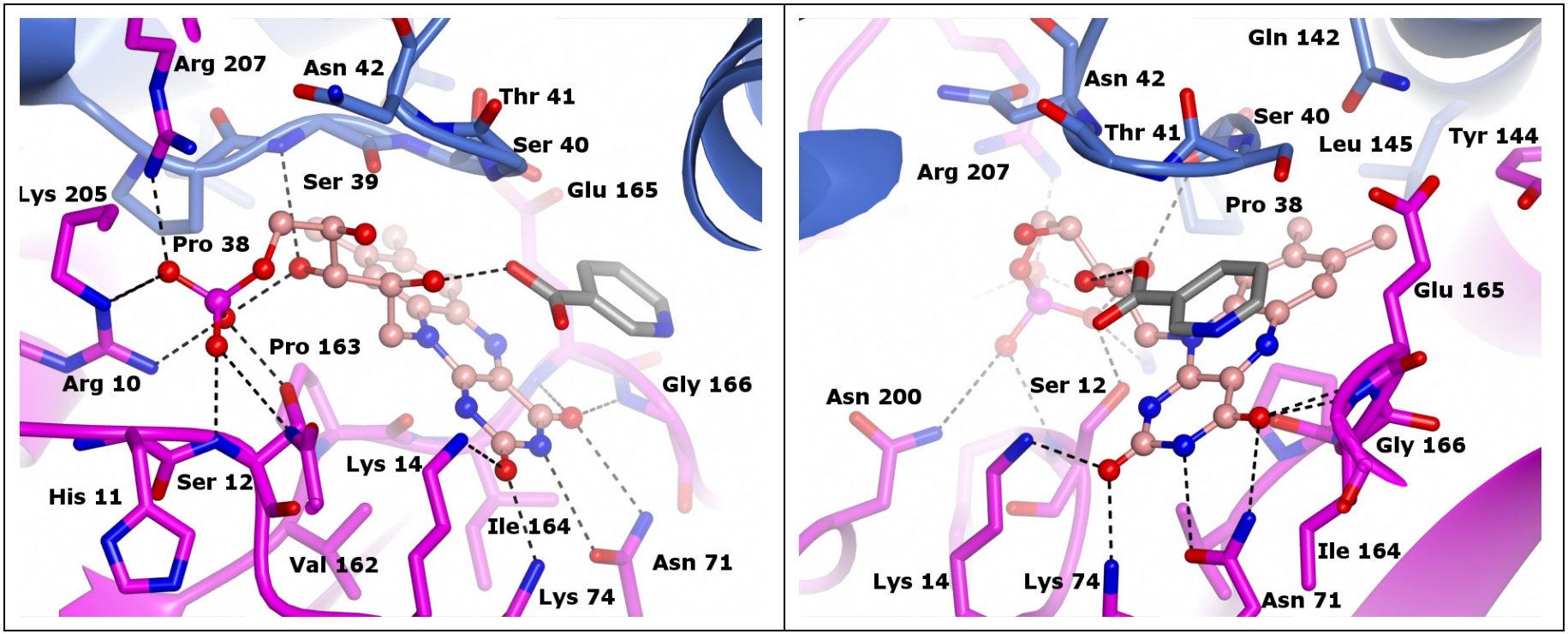


Figure 1-24: FMN binding site in NfsB, from the Nicotinate bound structure, pdb I.D. 1ICR (Lovering *et al.*, 2001). The FMN is shown as a pink ball and stick model, the nicotinate is shown in grey. All residues within 4.0 Å of the FMN are shown, coloured by chain so it can be seen that residues from both halves of the dimer form the binding site. Hydrogen bonds are shown as black dashed lines. Two images are separated by 90 °. Figures produced using CCP4 Molecular Graphics.

1.6.2.2.3 Ligand Binding in NfsB

Lovering *et al.* describe three crystal forms of NfsB with bound inhibitor nicotinate. (Lovering *et al.*, 2001) (Figure 1-21). Nicotinate is an analogue of the nicotinamide group of NAD(P)H and nicotinic acid has been shown to act as a competitive inhibitor for NADH (Race *et al.*, 2005). The nicotinate ligand was bound in the same orientation as benzoate was found to bind to the crystal structure of NR from *E. cloacae* (Haynes *et al.*, 2002). The crystal structure of NfsB has also been determined with the substrate, nitrofurazone (Figure 1-22), bound in the active site as well another with the inhibitor acetate (Race *et al.*, 2005). Nicotinate, acetate and nitrofurazone form one hydrogen bond to the O2* alcohol group of the FMN ribityl, with nicotinate and acetate otherwise bound through van der Waals contacts. The nitrofurazone was bound with the amide group over the catalytic N5 nitrogen of the FMN and the nitrofuran ring facing out towards the solvent, this may favour the electron transfer mechanism as discussed earlier (Section 1.6.1.1). The amide group hydrogen bonds to the side chain of Glu165, and the backbone amide of Thr41, while one oxygen of the nitro group hydrogen bonds to the backbone amides of Phe70 and Asn71.

Crystal structures of NfsB have also been published where CB1954 and the related dinitrobenzamide prodrug SN23862 are modelled to density inside the active site (Johansson *et al.*, 2003) (shown for CB1954 in Figure 1-25). The two non-crystallographically related active sites in the structure each have the prodrugs modelled in a different way. In one, the 2-nitro group of the prodrug (same for CB1954 and SN23862) is over the catalytic N5 nitrogen; in the other the 4-nitro group is. The molecule in the second active site is effectively flipped 180° relative to the

first, so that the opposite face is oriented towards the flavin. However there are doubts over these structures. For example the density is ambiguous and the molecular dynamics studies mentioned earlier (Section 1.6.1.1), used the binding shown by these structures as a model, but CB1954 came out of the active site in the studies and in two cases returned and reoriented itself on re-entry to have the amide over the N5 nitrogen (Christofferson and Wilkie, 2009). These molecular dynamics studies favoured a binding model based on the position of the amide from the nitrofurazone-bound NfsB structure (Race *et al.*, 2005). Furthermore, at the pH of the crystals (pH 4.6), the aziridine ring of CB1954 should ring open, and repeated attempts to reproduce these crystal soaks failed to remove the acetate from the NfsB active site. The exact structural mode of prodrug binding to NfsB is still unclear and further study is warranted.

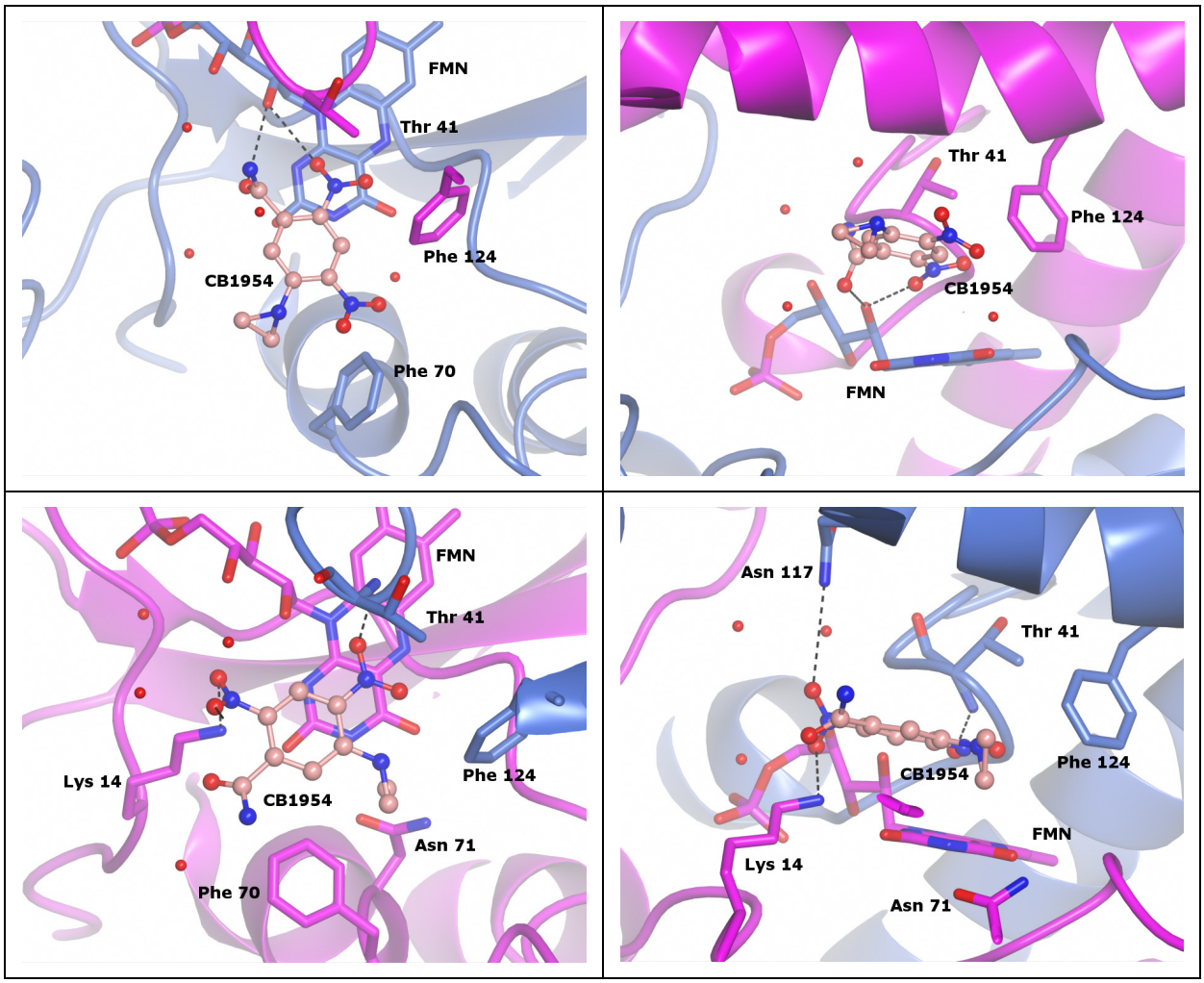


Figure 1-25: CB1954 (light pink) modelled in two active sites of NfsB by Johansson *et al.* (pdb I.D. 1IDT)(Johansson *et al.*, 2003). The black dashed lines represent H-bonds. The top two images are of the first active site and the bottom two of the second. Both sets of images are 90° apart. Each active site has a different nitro group above the flavin ring system; the first active site has the 2-nitro, the second the 4-nitro. Figures produced using CCP4 Molecular Graphics.

1.6.2.3 Mutants of NfsB

1.6.2.3.1 Generation of Mutants

The crystal structure of NfsB (Lovering *et al.*, 2001) identified residues in and around the active site that were most likely involved in substrate binding, and nine key residues were targeted for mutagenesis (Grove *et al.*, 2003). The nine residues targeted were Ser40, Thr41, Tyr68, Phe70, Asn71, Gly120, Phe124, Glu165 and Gly166, and

are shown in Figure 1-26. Libraries of random mutants were generated at each of these residues and the mutants were cloned into a bacteriophage λ vector, which was then used to infect NTR-deficient *E. coli* (UT5600 strain) to produce bacterial lysogens. These lysogens were then evaluated for their sensitivity to CB1954 by replica plating (Grove *et al.*, 2003). The only residue found to be essential for the catalysis of CB1954 reduction was Gly166, whilst no mutants with greater sensitivity to CB1954 were generated at the neighbouring Glu165. These residues were both identified earlier (Section 1.6.2.2.2) as being involved in binding the FMN, through hydrogen bonding by the backbone amides (Lovering *et al.*, 2001). Thus mutating Gly166 would be unfavourable as there would be no room for a side chain due to the flavin binding site, whilst Glu165 may be required for binding NAD(P)H. Mutations at Gly120 also provided no significant improvement but at all other sites mutants were found that improved sensitivity to CB1954. At Tyr68, Phe70 and Phe124 a diverse range of mutants were identified, and Phe124 showed improvement for 14 mutants. The most active of these were F124N and F124K. The F124K mutant was tested in SKOV-3 ovarian cancer cells and showed ~5-fold greater sensitization to CB1954 than WT (Grove *et al.*, 2003).

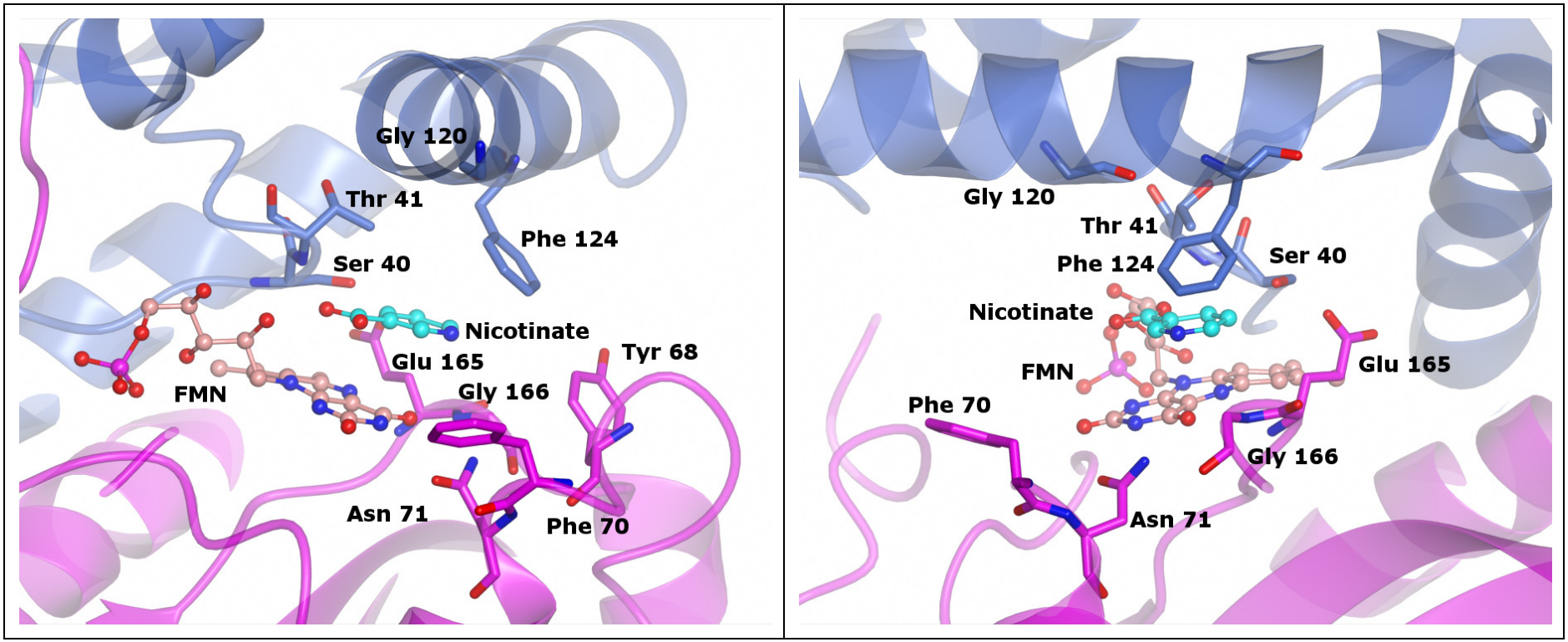


Figure 1-26: Location of the nine residues in the active site of NfsB selected for mutation by Grove *et al.* (Grove *et al.*, 2003). Image from the active site of the nicotinate-bound (shown in light blue) crystal structure (pdb I.D. 1ICR (Lovering *et al.*, 2001)), used by Grove *et al.* to identify the residues. The FMN prosthetic group is shown in light pink. The two images are 90° apart. Figures produced using CCP4 Molecular Graphics.

Single mutations at 6 of the 9 sites were combined in a subsequent study that generated 53 double mutant combinations (Jaberipour *et al.*, 2010). Lysogenised *E. coli* were replica plated onto media containing various concentrations of CB1954 and then assayed for their colony-forming ability. The 7 most promising mutants were then cloned into an adenovirus vector and tested for their ability to sensitize SKOV-3 cells to three prodrugs; CB1954, the dinitrobenzamide mustard SN23862, and the nitrobenzylphosphoramide prodrug LH7. The most active mutants were T41L/N71S and T41L/F170A for CB1954 (14-17-fold improvement over WT), T41L/F70A for SN23862 (4.8-fold improvement) and S40A/F124M for LH7 (1.7-fold). T41L/N71S was tested in two tumour xenograft models (SKOV-3 and human prostate carcinoma PC3 cells) and showed greater anti-tumour activity than WT NfsB.

A novel method of selecting successful mutants has been developed, exploiting CB1954 triggering the SOS response of *E. coli* (Guise *et al.*, 2007). DNA damage, such as the cross-linking caused by CB1954, activates the SOS response in *E. coli*, which activates the DNA repair mechanism (Sassanfar and Roberts, 1990). This occurs through the action of the RecA protein, which also promotes bacteriophage λ's transition into the lytic cycle through proteolytic cleavage of the cI transcriptional repressor protein (Roberts and Devoret, 1983). In bacterial lysogens, the phage DNA is incorporated into the bacteria and cI inhibits transition to the lytic cycle and so the prophage DNA is passed on to subsequent bacterial lysogen generations. Upon cI cleavage by RecA following DNA damage, this inhibition is lifted with the result that the cells enter the lytic cycle. The prophage is then excised, replicated and multiple phage particles are released upon lysis of the bacterial host cell (Hendrix *et al.*, 1983).

The induction of lytic cycle, with phage replica-plated, means that one can select for mutants with enhanced sensitivity to CB1954. The efficiency of selection was improved through repeated cycles of prodrug treatment of lysogens, phage harvest, and re-infection to generate more lysogens (Guise *et al.*, 2007).

This method was tested on randomised single mutations at either the N71 or the F124 sites, and then used to select the best mutants from a library with randomised mutations at three out of five sites (Guise *et al.*, 2007). All of these were mutated at F124, at either S40 or T41, and either F70 or N71. A lambda library was created of these mutants (a possible 1,048,576 nucleotide combinations), and the most active mutants were selected via 3 cycles of prophage selection at 100 μM CB1954, followed by 4 rounds at 50 μM and 3 rounds at 30 μM. Whereas before selection, ~ 90 % of lysogens were resistant at 400 μM CB1954, after selection, none could survive at just 25 μM. The NfsB genes generated from plaque-purified phage were sequenced to determine successful mutations. This selection method discovered one single mutant (F124N – previously identified (Grove *et al.*, 2003)), seventeen double mutants, 23 triple mutants and even 7 quadruple mutants (with mutations at other sites presumably incurred through errors in the PCR). The most active of these mutants as judged from the apparent IC_{50} of the lysogens was the triple mutant T41Q/N71S/F124T (IC_{50} = 3 μM, WT IC_{50} = 143 μM), a ~48 fold improvement. Interestingly the T41Q single mutant was worse than WT, and so the combination of mutations seen in this triple mutant shows cooperativity. Successful mutations at Thr41, Asn71 and Phe124, were used as the basis for another mutation library adding in mutations of Met127, His128 and Trp138 to all possible amino acids (unpublished work within this laboratory). Of

these the most successful mutant was T41Q/N71S/F124T/M127V, which showed a similar improvement in IC_{50} for CB1954 relative to WT as the triple mutant that included the first three mutations.

1.6.2.3.2 Analysis of Mutants

For most mutants, initially the IC_{50} for CB1954 (the concentration of CB1954 at which 50 % of cells survive) in *E. coli* lysogens was measured (Grove *et al.*, 2003, Guise *et al.*, 2007, Jaberipour *et al.*, 2010). The mutants with the lowest IC_{50} were then assessed in SKOV-3 cells, to confirm activity in 'target' cell lines, and most showed similar improvement to the *E. coli* lysogens over WT, although there were some notable exceptions. The most active single mutants, the double mutants N71S/F124K and N71S/T41L, and the triple mutant T41Q/N71S/F124T have also been purified and assessed *in vitro* for their ability to reduce CB1954 and nitrofurazone, where the specificity constant (k_{cat}/K_m) was compared as a measure of the activity of the mutant with CB1954 (Searle *et al.*, 2004, Race *et al.*, 2007, Jarrom *et al.*, 2009). The most active of these (T41L/N71S) had a ~100-fold higher k_{cat}/K_m than WT, and the next most active (T41Q/N71S/F124T) was ~45-fold higher. Mutants showing an improved k_{cat}/K_m for CB1954 showed improved kinetics for menadione but not for nitrofurazone. The WT global kinetic parameters for CB1954 were not measured as at high NAD(P)H concentration the K_{mapp} is higher than the solubility of CB1954. Since the plateau of the Michaelis-Menten curve is not seen the estimates of k_{cat} and K_m are poor, but it does give a good estimate of k_{cat}/K_m, which is estimated only on the, early, linear part of the Michaelis-Menten curve (Jarrom *et al.*, 2009). For a bi-bi 'ping-pong' mechanism, as the NAD(P)H concentration is raised the apparent k_{cat} and K_m

will increase until the NAD(P)H is saturating, but the k_{cat}/K_m will stay constant. All the mutants in this study appeared to show a greater specificity constant for CB1954 over nitrofurazone compared to WT, with the most improved being F124N and N71S/F124K showing ~5-fold improvement in specificity. Enhanced selectivity for the prodrug may be essential as it is known that NfsB has a broad substrate *in vivo* specificity and it is unknown how many substrates or inhibitors may interact with the enzyme. Some of the most active mutants for CB1954, have also been compared to WT for their ability to sensitize SKOV-3 cells to the alternative prodrugs SN23862 and LH7 (Jaberipour *et al.*, 2010) (see Section 1.5.2)

The reaction products of CB1954 reduction by NfsB and some of its mutants were assayed by HPLC (Race *et al.*, 2007, Jarrom *et al.*, 2009). The only deviation from the 1:1 ratio of the 2-hydroxylamine and 4-hydroxylamine products seen with WT was witnessed with the two mutants containing T41L, which seems to favour almost exclusive reduction of the 4-nitro group over the 2-nitro group. The replacement of threonine with leucine may have an effect on selectivity as the hydrophobic leucine would be more likely to form favourable interactions with the aziridine group of CB1954 and not the amide or nitro groups.

1.6.2.3.3 NfsB Mutant Structures

The crystal structures of four single mutants, T41L, N71S, F124K and F124N and one double mutant, N71S/F124K, have been determined, all bound to the substrate

analogue nicotinic acid (Race *et al.*, 2007). They allow some explanation of the improvements seen in activity for CB1954. There were no major changes in the overall structure, and the differences seen were in the neighbourhood of the mutation sites. For the T41L mutant, the backbone atoms are in identical positions to WT and hydrogen-bond the bound nicotinate through the backbone nitrogen. The leucine side chain in part overlays the threonine, essentially replacing the threonine alcohol group with a methyl group, but the leucine is larger and thus causes a slight change in the position of the nicotinate in the active site and the F124 side chain. The N71S mutation removes two direct hydrogen bonds from the FMN cofactor to the protein and replaces them with a hydrogen bond to water which is then in turn hydrogen bonded to the FMN (Figure 1-27). Serine was the only residue that showed improvement over WT at position N71 in saturation mutagenesis (Grove *et al.*, 2003). The F124N mutation moves the sidechain from a hydrophobic pocket for the phenylalanine to a more polar pocket for the asparagine. The lysine side chain of the F124K mutation is seen in multiple conformations with some being 'phenylalanine like' and others being 'asparagine-like'. In the most improved mutants, F124K and F124N, this movement of the side chain away from the active site and into a polar pocket, creates room for the relatively large CB1954 to move into the active site. Most mutations at F124 have been shown to be favourable in improving specificity for CB1954; this is most likely due to the sterical removal of a large group from around the active site.

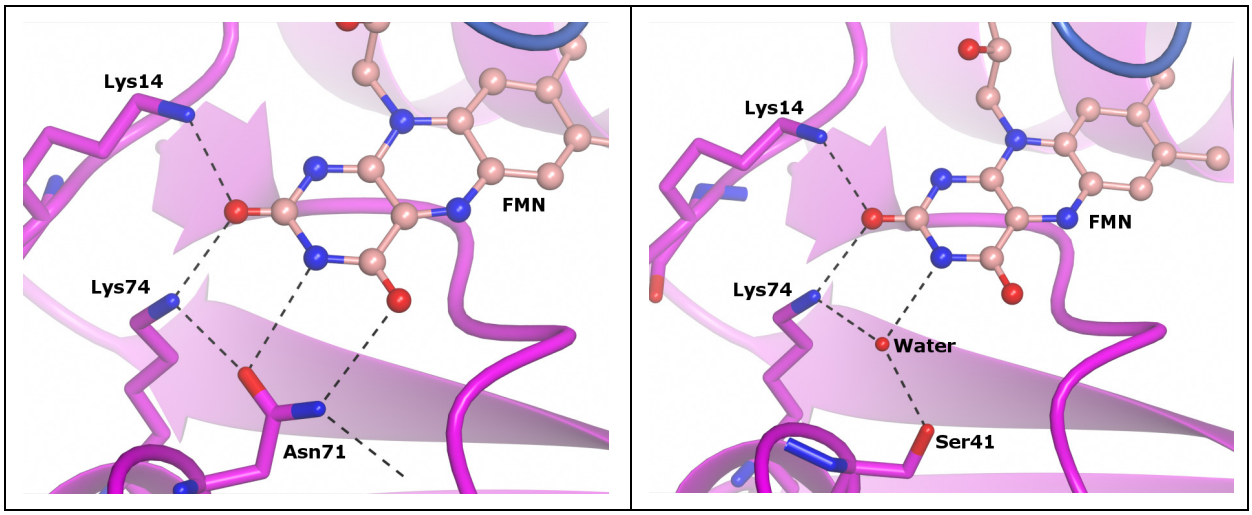


Figure 1-27: Effect of the N71S mutation in NfsB. The left-hand image is WT NfsB and the right-hand image is N71S. H-bonds are represented by black dashed lines. The direct hydrogen-bond to the FMN in WT is replaced by one through a solvent water. Figures produced using CCP4 Molecular Graphics.

1.6.2.3.4 Mutants containing Unnatural Amino Acids

The F124 residue has also been targeted for substitution with unnatural amino acids. These mutants were then tested for their activity with CB1954 *in vitro* (Jackson *et al.*, 2006). The synthetic amino acids are incorporated by exploiting the natural translation mechanisms of *E. coli*. The system requires the incorporation of two plasmids. Firstly a pBAD plasmid which incorporates the *nfsB* gene, but with an N-terminal 6 His-tag and the codon encoding the Phe124 residue mutated to the TAG stop codon. The plasmid also encodes an enterokinase cleavage site, between the His-tag and the N-terminus. Secondly a pDule plasmid is used that incorporates the gene encoding an orthogonal aminoacyl tRNA synthetase specific to the unnatural amino acid to be incorporated. The tRNA will incorporate the appropriate synthetic amino acid in response to amber (stop) codon UAG in mRNA. A different pDule plasmid is required for each different synthetic amino acid to be incorporated. This mechanism allows 100% incorporation of the unnatural amino acid and also high yields of protein. The

most active F124 unnatural amino acid substitution tested (pNF - *para*-nitrophenylalanine) showed a >30-fold improvement in k_{cat}/K_m over WT, and >2.3-fold better than the best of the natural amino acid substitutions (F124K). Other successful unnatural amino acid mutations included pAF (*para*-aminophenylalanine) and NapA (napthylalanine). The mutants were also tested for activity with LH7. Much smaller increases in activity were observed compared to WT, but the pNF was still the most active.

1.6.3 Properties of NfsA

The initial purification of NfsB based on its reduction of CB1954 used NADH as the assay cofactor (Anlezark *et al.*, 1992), and this may explain why NfsA was overlooked as a PAE since NfsA has a preference for NADPH as a cofactor (Zenno *et al.*, 1996a). A study in 2006 showed that NfsA can reduce CB1954 (Barak *et al.*, 2006), and a more recent study has compared NfsA to NfsB for potential use in suicide gene therapy (Vass *et al.*, 2009).

NfsA is a 27 kDa homodimer containing 240 amino acid residues, 23 more than NfsB. Alignment of the sequences shows only 28 identical and 32 conserved residues, and antibodies that are specific for NfsB do not cross react with the NfsA protein (Vass *et al.*, 2009). Whereas there is little sequence homology between the two *E. coli* nitroreductases, NfsA and NfsB, there is some structural homology (discussed in Section 6.2.1). *NfsA* expression is controlled by the *soxRS* regulon and has been shown to be induced by paraquat (Liochev *et al.*, 1999), unlike *nfsB* which is controlled by

the MarA regulon (Barbosa and Levy, 2002). Induction with 0.2 mM paraquat caused nitroreductase positive bacteria to display a 4.5 times greater nitroreductase activity with nitrofurantoin (Lightfoot *et al.*, 2000). NfsA may be expressed to control damage caused by oxidative stress, in common with other members of the *soxRS* regulon, by minimising the redox cycling from single electron reduction of nitroaromatics, quinones and dyes.

Overexpression of NfsA in *E. coli* lysogens caused greater sensitivity to CB1954 than for NfsB (>2-fold improvement) (Vass *et al.*, 2009). When expressed in SKOV-3 cells using a comparable adenovirus vector, NfsA showed between 3.5- and 8.1-fold lower IC_{50} than NfsB (in three experiments) (although antibodies were not available to confirm similar levels of expression). HPLC studies revealed that NfsA preferentially reduces the 2-nitro group of CB1954 over the 4-nitro, unlike NfsB which produces equal proportions of each hydroxylamine product. The 2-hydroxylamine has a pronounced bystander effect (Helsby *et al.*, 2004), and this is borne out by experiments in SKOV-3 cells when a minority of adenovirus infected cells are mixed with uninfected cells, NfsA shows a greater bystander effect than NfsB. NfsA was also compared for its ability to sensitise SKOV-3 cells to the dinitrobenzamide mustard prodrugs SN24927, SN26209 and SN27653 (Figure 1-12). All of the dinitrobenzamide prodrugs require activation of the nitro *para* to the mustard moiety, this is equivalent to the 2-nitro group of CB1954 and so NfsA may be expected to be more effective at reducing these prodrugs. NfsA-infected SKOV-3 cells showed 1.8- to 2.8-fold greater sensitivity to these prodrugs than NfsB infected SKOV-3 cells.

1.6.3.1 Kinetics of NfsA

Similarly to NfsB, NfsA has a broad substrate range that includes a variety of quinones and nitroaromatics (Zenno *et al.*, 1996a). NfsA shows a strong preference for NADPH over NADH as a cofactor *in vitro*, unlike NfsB which has a very slight preference for NADH (Vass *et al.*, 2009). With its preferred cofactor NfsA shows a ~25-fold higher k_{cat}/K_m (0.15 ± 0.01 μM^{-1} s^{-1}) when compared to NfsB (0.0058 ± 0.0004 μM^{-1} s^{-1}). Though only a ~18-fold improvement was seen when the small amount of DMSO used to solubilise the CB1954 is taken into account. With NADPH provided as a cofactor, purified NfsA is ~20-fold better at killing SKOV-3 cells than NfsB. This was a higher improvement than was suggested from the *in vivo* studies in cancer cell lines; this may be due to the available levels of NADH and NADPH in the cell or due to variable inhibition by small molecules inside the cell. In many tumours glycolysis is upregulated, the Warburg effect, which leads to an increase in the cellular NADH concentration, and indirectly to an increase in NADPH concentration. In order to fully understand the relative merits of the enzymes the concentrations in the target tissues would have to be assessed.

1.6.3.2 Structure of NfsA

1.6.3.2.1 Overall Structure of NfsA

The structure of the NfsA protein has been previously determined to 1.7 Å resolution (Kobori *et al.*, 2001). It shows some similarities with NfsB which are discussed in greater detail in Section 6.2.1. It is also a dimer with two FMN-containing active sites and is shown in Figure 1-28. It contains 11 α-helices and 5 β-sheets, with a core of one

large central helix and 4 stranded antiparallel β-sheets surrounded mostly by α-helices. The protein is larger than NfsB (240 versus 217 amino acids) and has a larger dimer interface, with 66 hydrogen-bonding and 609 non-bonding interactions involving approximately 98 residues from each monomer (PDEsum - (Laskowski, 2009)).

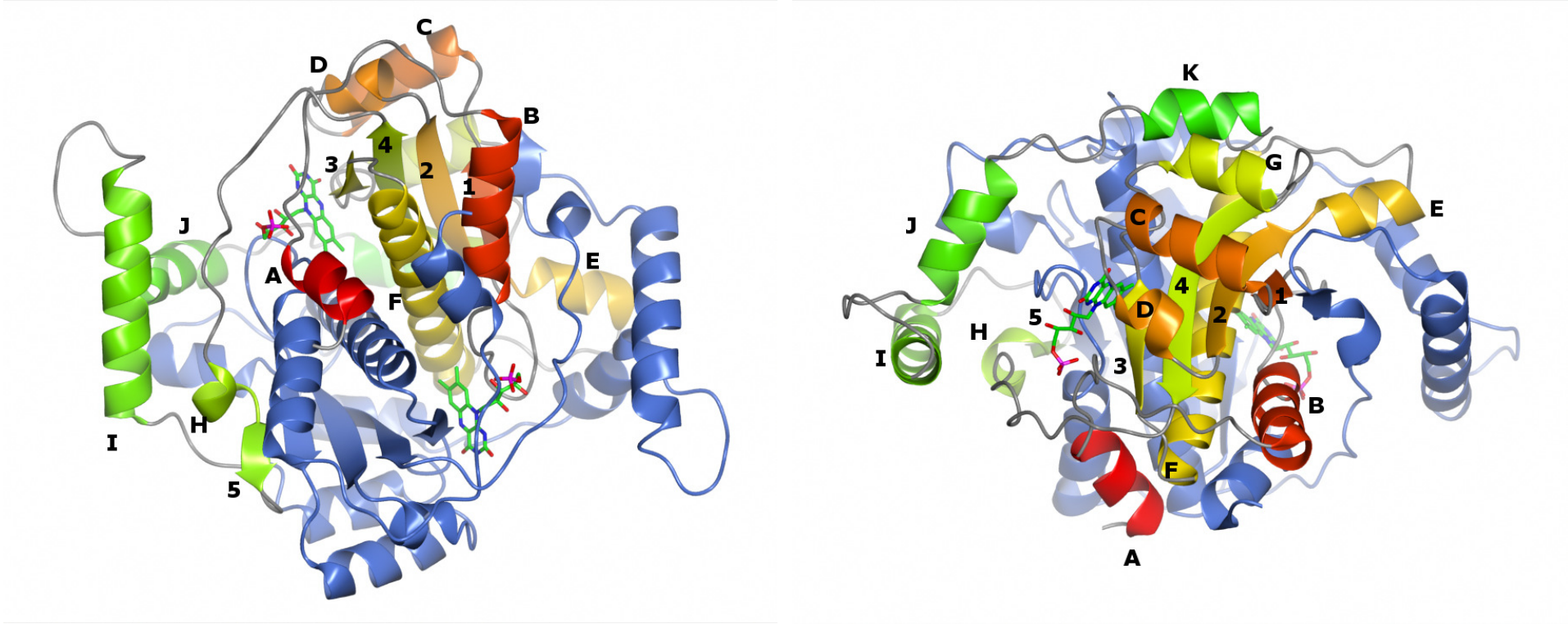


Figure 1-28: The crystal structure of the NfsA dimer (pdb I.D. 1F5V (Kobori *et al.*, 2001)). The two images are separated by 90°. A ribbon diagram of the protein is shown with one half of the dimer coloured in blue and the other by secondary structure succession. The secondary structure elements are labelled from the N-terminal, A-K for α-helices and 1-5 for β-sheets. The FMN prosthetic group is shown as green sticks and shows the position of the two active sites. Figures produced using CCP4 Molecular Graphics.

1.6.3.2.2 FMN Binding Site in NfsA

Similarly to NfsB, NfsA has two active sites found at the dimer interface (indicated by the green FMN models in Figure 1-28). The FMN binding site in NfsA is detailed in Figure 1-29. Similarly to NfsB, the flavin is bound primarily through hydrogen bonds to the ribityl and hydrophobic interactions with the ring system. There are 7 hydrogen-bonds to protein side chains and 3 to backbone atoms.

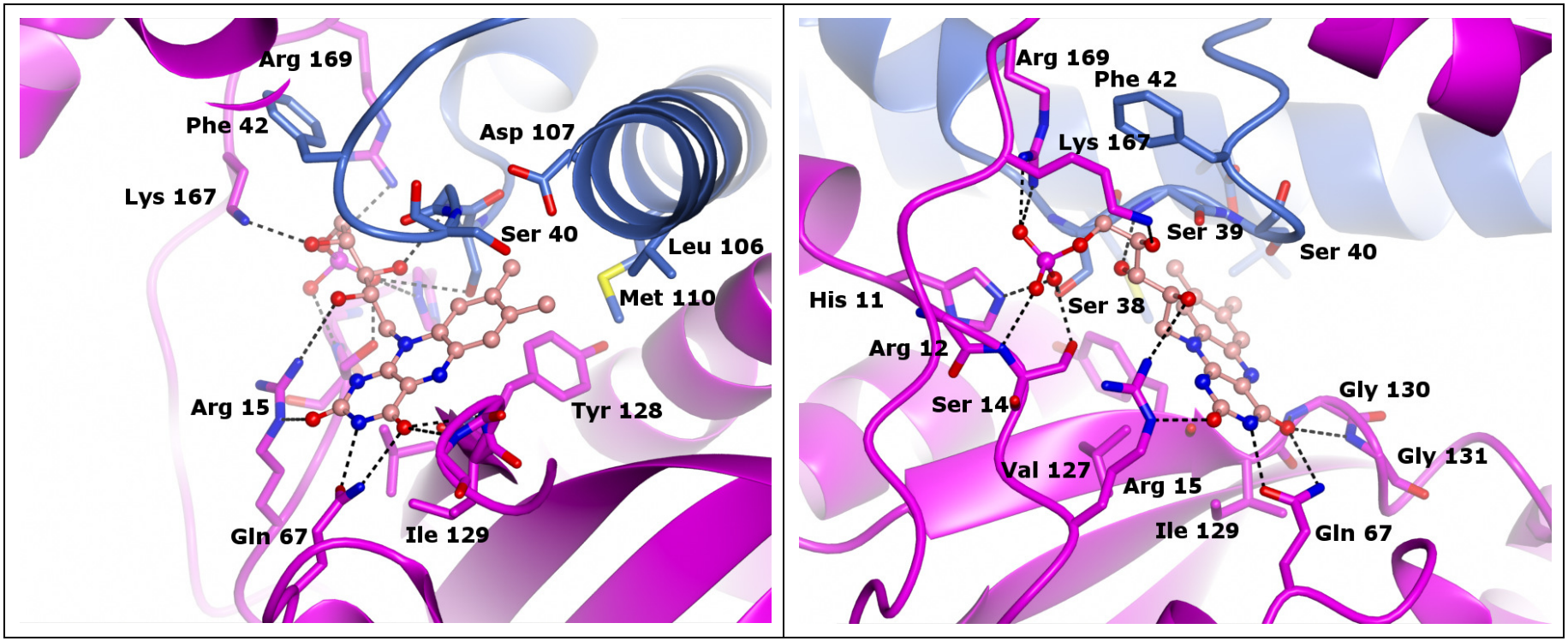


Figure 1-29: FMN binding site in NfsA. The Structure shown is the published Pdb I.D. 1F5V (Kobori *et al.*, 2001). The FMN is shown in pink, and the protein coloured by chain, to distinguish the dimer halves. H-bonds are represented by black dashed lines. Residues within 4.5 Å of the FMN are shown. The two images are separated by 90 °. Figures produced using CCP4 Molecular Graphics.

1.7 Aims of this project

This study will increase our understanding of the reduction of CB1954 by NfsA and NfsB. X-ray crystallography will be used to elucidate the structures of the NfsB mutants, and give insights into how the mutations improve CB1954 binding on a molecular level, building on studies with previous single and double mutants (Race *et al.*, 2007). Complementary kinetic studies will elucidate the effect the mutations have had on the various reaction steps. Finally crystallography will be used to understand the binding of ligands including substrates to NfsA, with the aim of identifying residues important in ligand binding possibly to guide future mutagenesis of the enzyme in a similar manner to that performed with NfsB (Grove *et al.*, 2003), and once again complementary kinetic studies will aid our understanding of the reactions taking place.

2. Materials and Methods

2.1 Materials

All reagents were purchased from Sigma Aldrich (Poole, UK) with the exceptions mentioned below. PEG 3000, 4000, and 8000 and the cytochrome c were purchased from Fluka Analytical (St. Gallen, Switzerland). CB1954 was kindly provided by ML Laboratories, Keele, UK. The other dinitrobenzamide (SN-) prodrugs were kindly provided by Dr. A. Patterson, Dr. J. Smaill and Prof. W. Wilson, Auckland Cancer Society Research Centre, The University of Auckland, Auckland, New Zealand. The unnatural amino acids, L-4-nitrophenylalanine (pNF), L-4-aminophenylalanine (pAF) and L-napthylalanine were purchased from PepTech Corp. (Burlington, Massachusetts, USA).

2.2 Production and Isolation of NfsB and NfsB Mutants

2.2.1 Growth of NfsB and Mutant Proteins

The Wild-type NfsB (Lovering *et al.*, 2001) and the T41L/N71S (Jarrom *et al.*, 2009) and T41Q/N71S/F124T (Guise *et al.*, 2007) mutants were made previously, as described. The T41Q/N71S/F124T/M127V mutant was identified and made using the lambda selection system described for the triple mutant. Successful mutations at Ser40, Thr41, Phe70, Asn71 and Phe124 were combined with mutations of Met127, His128 and Trp138 to all amino acids (unpublished work by Simon Vass). The mutant *nfsB* genes were cloned into pET11c plasmids under control of T7 RNA polymerase, which were transformed into *E. coli* BL21 (λDE3) cells. Colonies were grown and

selected for their resistance to kanamycin. Cells were cultured in Luria Broth (LB) at 37°C and grown until the scattering of light at 550 nm (OD_{550}) of broth solution was 0.7 - 0.9 units, at which point IPTG was added to a final concentration of 1 mM to induce the transcription of the T7 RNA polymerase in the lambda lysogen and hence the *nfsB* gene on the pET plasmid, and the flasks were incubated at 37°C for a further 4 to 6 hours. The induced cultures were centrifuged at 17,700 x g for 15 min to pellet the cells. The pellet was resuspended and washed in a minimal amount of STE buffer (10 mM Tris-HCl pH 7.0, 1 mM EDTA, 100 mM NaCl) and centrifuged again at 17,695 x g for 15 min.

2.2.2 Purification of NfsB and NfsB Mutants

NfsB and the natural amino acid NfsB mutants were purified following the method modified from that described by Anlezark *et al.* (Anlezark *et al.*, 1992). Cell pellets were resuspended in a minimal volume of 10 mM Tris pH 7.0, 10 µg/ml DNaseI, 10 µg/ml RNaseA, 10 µg/ml PMSF, 0.1% iso-propanol, 20 µM FMN, 10 mM $MgSO_4$, 1 mM $MnCl_2$. The cells were lysed by sonication with a 'Heat Systems Sonicator', at power level 5, using three times a programme of 15 second bursts with 20 second rests for a total time of one minute, and then centrifuged at 33,000 x g for 30 min. The supernatant was collected and ammonium sulphate was added to 0.5 M. The solution turned a cloudy golden yellow, and this solution was then filtered.

The solution was loaded onto a Phenyl-Sepharose hydrophobic interaction chromatography column pre-equilibrated with 10 mM $NaPO_4$ pH 7.0, 1.0 M

$(NH_4)_2SO_4$, 20 µM FMN, 10 µg/ml PMSF, 0.1% iso-propanol. A wide golden yellow band appeared at the top of the column, due to the nitroreductase-bound FMN (a good indicator of the proteins location). The column was then washed with 3.5-4 times column-volume of the above buffer. The protein was eluted using 10 mM Tris pH 7.5, 20 µM FMN, 10 µg/ml PMSF, 0.1% iso-propanol. Fractions were assessed for protein content as described in Section 2.2.3.

Selected fractions were pooled and dialysed into 20 mM Tris pH 7.5, 20 µM FMN, 10 µg/ml PMSF, and 0.1% iso-propanol and then loaded onto a Q-Sepharose ion-exchange chromatography column pre-equilibrated in the same buffer. The column was washed with 3.5-4 times column-volume of buffer. The column was eluted with buffer containing 100 mM NaCl to 500 mM NaCl, followed by buffer with 500 mM NaCl. The NfsB content of samples was assessed and fractions were pooled on the basis of specific activity and concentrated as described in Section 2.2.3.

The NfsB quadruple mutant (T41Q/N71S/F124T/M127V) was expressed in *E. coli* and purified by Alex Gavin, a final-year project student.

2.2.3 Assessment of Enzyme Purity and Concentration

2.2.3.1 SDS-PAGE

Purity of fractions was assessed by Coomassie brilliant-blue stained SDS–PAGE carried out as described previously (Laemmli, 1970), using Biorad mini-protean

apparatus (Biorad, Hemel Hempsted, UK). All gels contained 12% acrylamide (30:0.8% Acrylamide:Bis-acrylamide – pre-mixed), and at least one lane containing Biorad Broad-Range molecular weight markers.

2.2.3.2 Protein Concentration Determination

The concentration of the protein was estimated by measuring the absorption of light at 280 nm, using an extinction coefficient of 43,000 M^{-1} cm^{-1} for NfsB and 52,000 M^{-1} cm^{-1} for NfsA (purification detailed in Section 2.4), and by Bradford Assay, using Bovine Serum Albumin (BSA) as a standard (Bradford, 1976). For the Bradford assay, 500 µl of Bradford reagent was added to a total volume of 1 ml. The absorbance at 550 nm was then compared to a standard curve created using known concentrations of BSA up to 10 µg/ml to assess the protein concentration.

2.2.3.3 Nitroreductase Activity Assays

The nitroreductase activity was assessed via a menadione-linked cytochrome c activity assay, whereby the reduction of menadione by the nitroreductase is linked to the reduction of cytochrome c by the product of menadione reduction, menadiol (Figure 2-1). Typically 1 ml solutions were prepared directly in a cuvette containing 10 mM Tris-HCl pH 7.0, 200 µM menadione, 80 µM cytochrome c (from Horse Heart), 100 µM NADPH, 2% DMSO. The reaction was started by the addition of the enzyme from the fraction to be tested, typically 0.5-2 µl. After mixing the solutions, the rate of reaction was measured from the change in absorbance of cytochrome c at 550 nm over a period of two minutes at 25°C. Rates were calculated assuming a molar absorbance

for cytochrome c of 21,100 M^{-1} cm^{-1} at 550 nm and dividing by 2, as two moles of cytochrome c are reduced for each mole of menadione. It has been shown that, at this concentration of cytochrome c and low NfsB concentrations, the rate of reaction is the same as the rate for the reduction of menadione alone (Jarrom *et al.*, 2009).

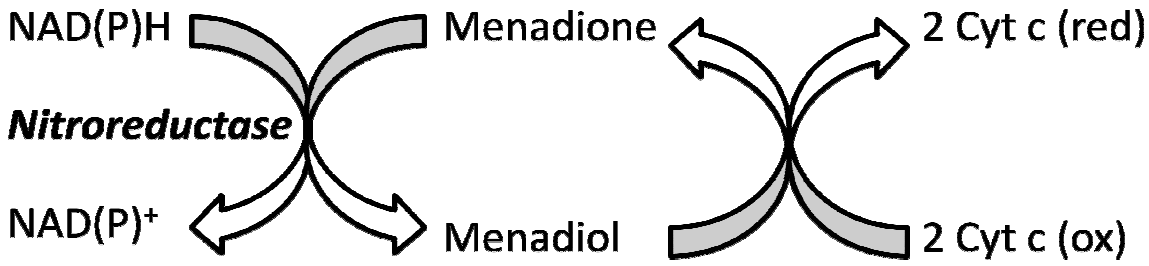


Figure 2-1: Reaction scheme for Menadione-linked Cytochrome C nitroreductase activity assay.

2.2.3.4 Pooling and Concentration of Enzyme Fractions

Conductivity of fractions from the columns was measured using a conductivity meter (Radiometer Copenhagen, Denmark). Specific activities were calculated by dividing the activity (AU min^{-1} μg^{-1} enzyme) by the protein concentration and fractions with the highest specific activities were pooled and concentrated using Centricon® YM-10 microconcentrators (Millipore, Watford, UK).

2.3 Production and Isolation of His-tagged NfsB and Unnatural Amino Acid Mutants

2.3.1 Source of Plasmids and Cells

Professor Ryan Mehl (Oregon State University, USA) kindly provided the plasmids; pBAD-NTR, pBAD-NTR-124TAG and pDule-Nap, as well as *E. coli* DH10B cells transformed with pBAD-NTR-124TAG and either pDule-pAF or pDule-pNF. Vector maps of the generic pBAD and pDule plasmids are shown in Figure 2-2. The production of these plasmids is unpublished, but follows a method described by Hammill *et al.* (Hammill *et al.*, 2007) to produce mutants containing the amino acids described previously (Jackson *et al.*, 2006). The NfsB gene from the *E. coli* genome was cloned into the pBAD vector, producing pBAD-NTR, and then the 124 residue was mutated to a stop codon TAG, producing pBAD-NTR-124TAG. The cloning event removes the first codon and adds a large affinity tag to the N-terminal of the protein. The full sequence of the tag is shown in Table 2-1. This expresses tagged NfsB under the control of the pBAD promoter, which is induced by arabinose (Figure 2-2). The pDule plasmids encode the aminoacyl tRNA synthetase that adds the appropriate unnatural amino acid to the suppressor $tRNA_{CUA}$, which recognises the UAG sequence in mRNA (Figure 2-2).

ATG GGG GGT TCT CAT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG
Met Gly Gly Ser His His His His His His Gly Met Ala Ser Met

ACT GGT GGA CAG CAA ATG GGT CGG ACT CTG TAC GAC GAT GAC GAT
Thr Gly Gly Gln Gln Met Gly Arg Thr Leu Tyr Asp Asp Asp Asp

AAG GAT CGA TGG GGA TCC
Lys Asp Arg Trp Gly Ser

Table 2-1: Sequence of His-tag attached to NfsB in the pBAD-NTR vector. The DNA sequence is shown in italic, and the amino acid sequence in bold. The first underlined section is the His-tag, and the second the enterokinase recognition site. The tag is cleaved after the Lys, meaning the last 5 amino acids still remain attached to the protein in place of the N-terminal methionine.

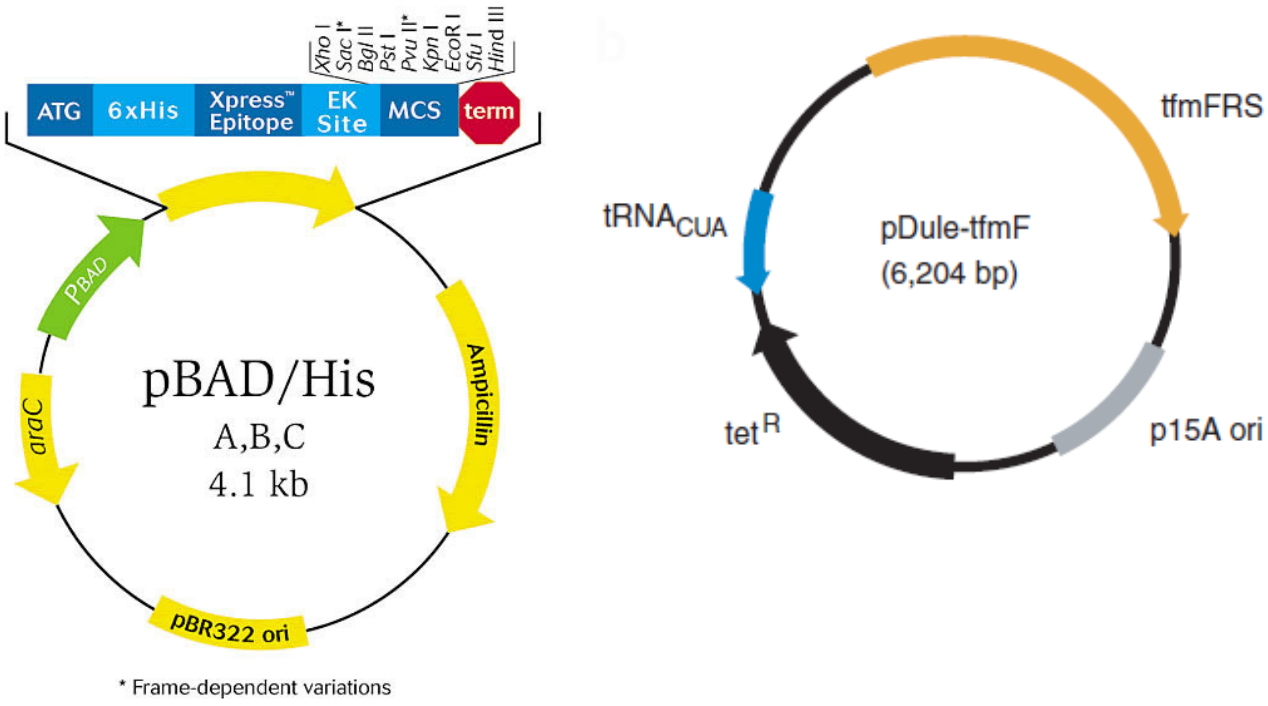


Figure 2-2: Vector map of the pBAD/His plasmid (Invitrogen) and the pDule-tfmF plasmid adapted from reference (Hammill *et al.*, 2007). The pBAD plasmid contains the inducible promoter and adds a 6xHis tag to the N-terminal of the protein added into the multiple cloning site (MCS). An enterokinase (EK) cleavage site is also added, allowing the His-tag to be cleaved. The pDule-tfmF plasmid is a plasmid designed to work in conjunction with the pBAD plasmid to incorporate the unnatural amino acid tfmF (trifluromethylphenylalanine). The pDule plasmids used in this study replace the aminoacyl tRNA synthetase tfmFRS, with one specific to the unnatural amino acids studied here.

The pBAD plasmids contain an ampicillin resistance gene and the pDule plasmid a tetracycline resistance gene. *E. coli* DH10B competent cells, produced by electroporation, were transformed with the pNAD-NTR-124TAG and pDule-Nap plasmids to provide cells capable of growing the NfsB Nap mutants with napthylphenylalanine at position 124. Cells transformed with pBAD encoding WT

NfsB were selected by growth in 100 µg/ml ampicillin, and the mutants were selected with additional 25 µg/ml tetracycline.

2.3.2 Growth and purification of His-tagged NfsB and Unnatural Amino Acid Mutants

The His-tagged NfsB was grown using an arabinose auto-induction media (Studier, 2005). This works by using a growth medium containing glucose and arabinose as the sources of energy. *E. coli* has a preference for glucose and so the cells will initially grow using the glucose. The *nfsB* gene is placed under the control of the L-arabinose or *ara* operon, which encodes the enzymes required for the catabolism of arabinose. Glucose acts as a repressor of the *ara* operon, and so while glucose is present induction of the operon and the *nfsB* gene is repressed, but in the absence of glucose *nfsB* is transcribed. The amount of glucose is sufficient to grow the cells to mid-log, and then the gene is transcribed.

Growth of *E. coli*-expressing NfsB and purification of the His-tagged NfsB and unnatural amino acids was performed according to the procedure of Hammill *et al.* (Hammill *et al.*, 2007). An overnight culture of the cells was grown in 50 ml L-Broth containing 10 g/L Tryptone, 5 g/L Yeast Extract and 10 g/L NaCl. 2.5 ml of the saturated culture was added to 500 ml of arabinose auto-induction media (Studier, 2005), containing 0.05% glucose and 0.05% arabinose as energy sources, along with amino acids, salts and trace metal ions. FMN was also added to 20 µM. The unnatural amino acid of interest was added to a concentration of 1 mM 30 minutes after

inoculation from an overnight culture in LB. The overnight and auto-induction cultures also contained 100 µg/ml ampicillin and, for the mutants, 25 µg/ml tetracycline. Cultures were shaken at 250 rpm and kept at 37°C. Growth was measured from turbidity at 600 nm and NfsB expression was monitored by SDS-PAGE (as per Section 2.2.3.1), 0.1 ml of 5 M Urea loading dye (also containing 0.04% Bromophenol Blue, 20% Glycerol, 2% β-mercaptoethanol, 0.1% SDS and 150 mM Tris pH 6.8) per OD unit of sample was added. The samples were boiled and analysed on SDS PAGE. When growth was complete cells were pelleted as described in Section 2.2.1.

Cell pellets were resuspended in a 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 20 µM FMN solution containing a single protease inhibitor tablet (EDTA free, Roche). This solution was sonicated and cell debris spun down as in Section 2.2.2. The supernatant was then loaded onto a Nickel NTA column pre-equilibrated in 50 mM sodium phosphate pH 8.0, 300 mM NaCl and 10 mM imidazole. The column was washed with three column-volumes of buffer. Protein was eluted using a gradient of increasing imidazole, 20 mM to 250 mM over four column volumes. Protein fractions were analysed for the presence of nitroreductase by Bradford assays and SDS PAGE, as in Section 2.2.3. The purest fractions were then pooled and dialysed into 20 mM sodium phosphate pH 7.0, 0.5 M ammonium sulphate and 20 µM FMN overnight.

Protein fractions were then loaded onto a Phenyl-Sepharose column pre-equilibrated in 20 mM sodium phosphate, 0.5 M ammonium sulphate and 20 µM FMN, as used

previously with non His-tagged NfsB (Lovering *et al.*, 2001) (Section 2.2.2). The column was washed with two column-volumes of buffer and protein fractions were eluted using 10 mM Tris pH 7.0, 20 µM FMN. Fractions were analysed as described in Section 2.2.3 and pooled on the basis of specific activity.

2.3.3 Cleavage of His-tag

The C-terminal His-tag was cleaved using a recombinant enterokinase (rEK - 50 U stock, Novagen, UK). Protein to be cleaved was dialysed into 20 mM Tris pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$. To test cleavage, protein was diluted to 0.2 µg/ml and enterokinase added to a final concentration of 0.02 - 0.1 U in 50 µl. Test samples were left at room temperature for 48 hours, with samples taken periodically, to which was added protein loading dye (for SDS-PAGE) and 10 µg/ml PMSF was added to the samples to inhibit further cleavage. Final cleavage was performed by adding rEK to ~1 mg/ml NfsB to 0.5 U in 500 µl, leaving at room temperature for 48 hours and then adding 10 µg/ml PMSF to prevent further cleavage.

2.4 Production and Isolation of NfsA

2.4.1 Growth of NfsA

The *nfsA* containing pET-24 plasmid pPS1341A1 was made by Dr Peter Searle (Vass *et al.*, 2009), and used to transform *E. coli* BL21 (λDE3) cells. Colonies were grown and selected for their resistance to kanamycin. Cells were grown and production of NfsA induced as described for NfsB in Section 2.2.1.

2.4.2 Purification of NfsA

Purification was performed by a method similar to that for NfsB (described in Section 2.2.2), with the exceptions noted here. After the cell debris had been pelleted, ammonium sulphate was added to 208 g/L for a 35% cut. The cloudy golden-yellow solution was then centrifuged at 7,500 x g for 10 minutes. The supernatant was removed and to it was added 236 g/L of ammonium sulphate for a 35-75 % cut. The thick cloudy golden yellow suspension was centrifuged at 7,500 x g for 10 minutes. The yellow pellets were redissolved in a minimal amount of 10 mM Tris-HCl pH 7.0, 1 mM EDTA 1.0 M $(NH_4)_2SO_4$, 20 µM FMN, 10 µg/ml PMSF, 0.1% iso-propanol prior to loading onto the pre-equilibrated Phenyl-Sepharose column. A three step elution process followed, all buffers containing 10 mM Tris-HCl pH 7.0, 20 µM FMN, 10 µg/ml PMSF, 0.1% iso-propanol. The column was washed first with two column-volumes of 0.8 M $(NH_4)_2SO_4$ buffer, then two column-volumes of 0.4 M $(NH_4)_2SO_4$ buffer, and finally with 0.2 M $(NH_4)_2SO_4$ buffer. The column was finally washed with a no salt buffer to remove the final traces of protein. The Q-Sepharose column was then used in an identical manner to that used for NfsB.

If further purity was required the pooled fractions were run down a Blue Trisacryl affinity chromatography column, and washed through with further sample buffer. The column was pre-equilibrated with 40 mM sodium phosphate pH 7.0, 20 µM FMN, 10 µg/ml PMSF, 0.1% iso-propanol and the samples were dialyzed into the same buffer

prior to loading. The protein was eluted with a gradient run from the no salt buffer to a 1.0 M NaCl buffer.

2.5 X-ray crystallography

2.5.1 Crystallization

All crystals in this study were grown using a vapour diffusion method. Either by a sitting drop method, using CombiClover 96-well crystallization plates (Emerald BioSystems, Bainbridge Island, WA, USA), or a hanging drop method, using MD3-11 XRL plates and CrystalClene^{TM} cover slips (both - Molecular Dimensions Limited, Newmarket, Suffolk, UK). The cover slips were sealed to the tray using high vacuum grease. Sitting drop trays were sealed using clear tape. 1 ml mother liquor was used in the reservoir of all trays. Equal volumes of protein solution and mother liquor were mixed together in the crystal drops. Volumes were varied from 1 µl of each to 4 µl of each. All trays were covered with foil (to prevent light damage as the FMN is light-sensitive) and stored at a constant temperature of 18°C. All chemical solutions for crystallization were sterile-filtered prior to use and additives/ligands were added at the appropriate pH. Crystal photos were taken ten levels of focus and combined using the CombineZP programme, these were then cropped and white balanced in the photo editor GIMP2.

2.5.1.1 Crystallization of NfsB mutants

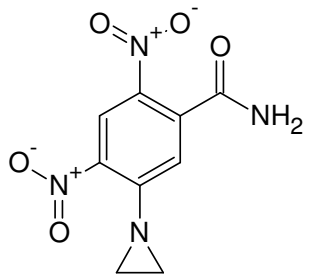
2.5.1.1.1 Crystals trays at pH 4.6

Initial screening was based on initial conditions reported by Lovering *et al.* (Lovering *et al.*, 2001). The hanging drop set-up was observed to produce higher quality crystals than the sitting drop method. Parameters varied included protein concentration, drop size, concentration of PEG 4,000 and concentration of sodium acetate. Crystals grew in a wide range of conditions, with reproducible crystals grown from 10-18% PEG 4,000 and 50-200 mM sodium acetate buffer pH 4.6. Protein concentrations that led to crystal growth were 8-12 mg/ml. Later conditions included up to 15% ethylene glycol, to aid in cryo-protection, or 15 mM nicotinic acid, to obtain ligand complexes, added to the mother liquor.

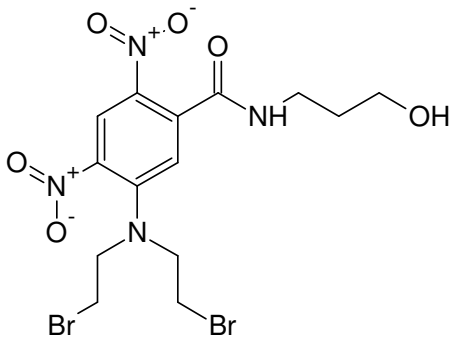
2.5.1.1.2 Crystal trays at pH 7.0

Crystals grown at pH 7.0 grew from conditions based upon those described by Parkinson *et al.* (Parkinson *et al.*, 2000). Crystals grew from mother liquor containing 15-22 % PEG 8,000 and 10-20 mM potassium phosphate pH 7.0. The first crystals obtained in these conditions were found to be very fragile plates and rarely survived attempts to cryo-protect by moving from one solution to another, and so to limit the stress on these crystals 15 % ethylene glycol or 15 % DMSO was added to the mother liquor when setting up the trays to cryo-protect and allow crystal mounting straight from the drop.

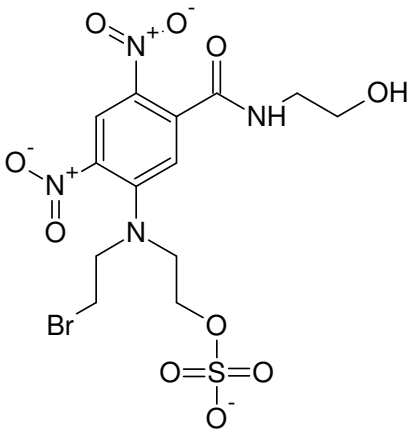
Attempting to obtain protein-ligand complexes, NfsB T41L/N71S and T41Q/N71S/F124T mutants were co-crystallized with up to 10 mM CB1954 (in a solvent of 2:7 NMP:PEG300) and up to 20 mM of the related dinitrobenzamide prodrugs SN27419, SN28099, SN27686 (Singleton *et al.*, 2007), S27683 and SN29222 (all in acetonitrile) (Figure 2-3). To save material, the prodrugs were added to the protein solution prior to mixing with the mother liquor in the crystal tray.



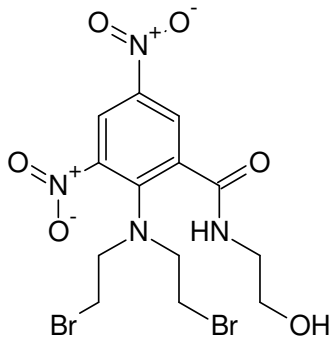
CB1954



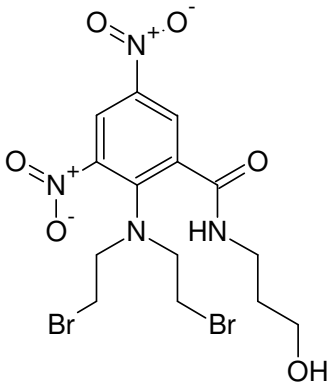
SN27419



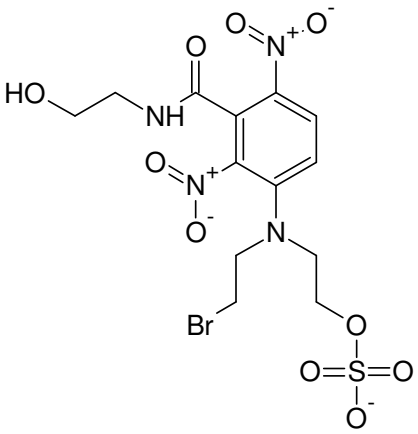
SN28099



SN27686



SN27683



SN29222

Figure 2-3: Dinitrobenzamide prodrugs used in crystallization studies. Figure produced using ChemSketch (ACD/labs).

2.5.1.2 Crystallization of NfsB Unnatural Amino Acid Mutants

Crystals of NfsB unnatural amino acid mutants were grown under similar conditions to the previous NfsB crystals at pH 4.6 (Section 2.5.1.1.1). For anything other than thin needles to form the His-tag had to be cleaved prior to crystallization as described in Section 2.3.3. Crystals grew from drops containing 10-15 mg/ml protein. Crystals of the pNF mutant were grown by Sophie Englander and of the pAF mutant by Laura Goodman (both undergraduate project students).

2.5.1.3 Crystallization of *NfsA*

Purified NfsA was concentrated to between 10-16 mg/ml (concentration determined by Bradford Assay as before), and then dialysed into 100 mM Imidazole. Some protein solutions were dialysed into buffer also containing 10 mM NaCl and/or 20 μM FMN. Conditions varied included drop size, precipitant concentration, buffer concentration and ratio of protein to mother liquor in drop. Protein concentration was varied on many trays, from approximately 5-16 mg/ml, by dilution into 100 mM Imidazole pH 7, before mixing with mother liquor. Crystals grew from mother liquor for NfsA crystallization contained 50-200 mM Imidazole pH 7 as a buffer and 18-34% PEG 3,350 (Sigma Aldrich, Poole, UK) or PEG 3,000 (Fluka Analytical, St. Gallen, Switzerland) as a precipitant.

2.5.2 Ligand Soaks

With NfsB crystals grown in acetate buffer, soaking in new ligands had only proven possible in the past by soaking the acetate-grown crystals in 100 mM sodium citrate

pH 4.6 along with the ligand of interest (Race *et al.*, 2005). Unsuccessful attempts were made to use this method to soak in substrates including menadione and nitrofurazone (up to 25 mM of each), so the decision was made to include no ligand and see the effect of citrate alone. Soaking was not possible with the fragile crystals of NfsB grown at neutral pH. With NfsA, it was only once the PEG used as a precipitant was changed from PEG 3,350 (Sigma) to PEG 3,000 (Fluka) and the cryo-protectant switched from ethylene glycol to DMSO that soaking proved effective. Then soaking with Hydroquinone (25 mM) and Benzoquinone (10 mM) in 5% steps gave rise to sufficient density in the activity site.

2.5.3 Cryo-protection and Crystal Mounting

To cryo-protect, the crystals were soaked in mother liquor containing either ethylene glycol or DMSO, increasing concentration in incremental steps of 5% for 1-10 minutes each to minimise osmotic shock, to a total concentration of 15% for ethylene glycol and 20% for DMSO. With NfsA crystals for each of these steps it was necessary to lower the concentration of the PEG precipitant by 2% (e.g. from 18% to 16%) alongside each incremental increase in cryo-protectant to prevent the crystals from dehydrating (when this occurred crystals showed a different unit cell and the structures could not be solved). The crystals were then flash-cooled in liquid nitrogen.

2.5.4 Data collection, processing and refinement

2.5.4.1 X-ray Sources

X-ray diffraction data sets were collected on either the University of Birmingham's home source (Rigaku 007HF generator with a Saturn CCD detector mounted on a 4-circle kappa goniometer) or at the synchrotrons of the European Synchrotron Radiation Facility (ESRF), Grenoble, France or Diamond Light Source, Oxfordshire, UK. The home source was frequently used to screen crystals before taking them to the synchrotron. The beamline will be specified in the results chapters for datasets collected at ESRF and Diamond.

2.5.4.2 Software used for data processing and refinement

Diffraction images were indexed, integrated and processed using MOSFLM (Leslie and Powell, 2007), iMOSFLM (Battye *et al.*) or XDS (Kabsch). Datasets were combined and scaled using POINTLESS and SCALA (Collaborative Computational Project, 1994) and data quality was assessed using XTRIAGE (Adams *et al.*, 2010). All structures shown here were solved by molecular replacement with PHASER (McCoy *et al.*, 2007), using published structures as starting/search models, pdb entry 1F5V for NfsA (Kobori *et al.*, 2001) and pdb entry 1ICR for NfsB (Lovering *et al.*, 2001). The NfsB search model was modified to replace the residues that were mutated with alanines so the density for the mutation sites could be judged with a less biased electron density map. An alternative search model was used for structures in a new space group, using only chain A of the 1ICR structure. Structures were refined using REFMAC5 (Murshudov *et al.*, 1997, Collaborative Computational Project, 1994) and

PHENIX (Adams *et al.*, 2010). For structures with high resolution (better than 1.1 Å) hydrogen atoms were added to the model prior to refinement using phenix.reduce (Word *et al.*, 1999). Models were built and modified using Coot (Emsley and Cowtan, 2004). Final models were validated using MOLPROBITY (Chen *et al.*, 2010) and POLYGON (Urzhumtseva *et al.*, 2009).

2.6 Enzyme Kinetics

All data sets were fitted to the stated equations by nonlinear regression with equal weighting of all points using SIGMAPLOT 12 (Systat Software Inc., Richmond, CA, USA). Statistical analysis was performed by means of the t-test, with t values giving the ratio of the regression coefficient and its standard error and P values giving the probably of the t value occurring by chance. For reliable data statistics, high t values (> 5) and low P values (< 0.05) are desirable.

2.6.1 Steady-State Enzyme Kinetics

2.6.1.1 Data collection and analysis

Experiments were performed using a Kontron Uvikon 923 spectrophotometer, using cuvettes with a 5 mm or 10 mm path length. Solutions were prepared directly in the cuvette and mixed by inversion.

2.6.1.2 Experimental conditions

All experiments were performed in 10 mM Tris-HCl pH 7.0, 50 mM NaCl and at 25°C. Reactions were initiated by the addition of a small quantity of enzyme (~10 nM). Nitrofurazone, nitrofurantoin and menadione were all dissolved in 90% DMSO, 10 mM Tris-HCl pH 7.0, while CB1954 was in a 2:7 mixture of NMP:PEG 300. Kinetic experiments involving menadione, NFN and NFZ all included final concentrations of 4.5 % DMSO. Experiments with CB1954 had final concentrations of 1.1 % NMP and 3.9 % PEG 300. Concentration ranges for substrate and inhibitors are detailed in the kinetics parameters tables for each experiment. Substrate concentration was either measured by monitoring the extinction coefficient as detailed in Table 2-2 or by weight. The wavelengths the reactions were followed at, the extinction coefficients for the reaction and the changes monitored are detailed in Table 2-3.

Substrate	Wavelength Monitored λ (nm)	ε at λ (M^{-1} cm^{-1})	Reference
NAD(P)H	340	6,200	(Galante and Hatefi, 1978)
CB1954	327	11,600	(Knox *et al.*, 1993a)
NFN	373	18,481	*Merck Index*
NFZ	400	12,960	(Peterson *et al.*, 1979)
Quinone	246	25,420	Unpublished work

Table 2-2: Extinction coefficients of substrates used in this study.

Substrate	Change Monitored	$\lambda_{reaction}$ (nm)	Change in ε at $\lambda_{reaction}$ (M^{-1} cm^{-1})
CB1954	2-/4-hydroxylamine formation	420	1,200
Menadione	Cytochrome c reduction	550	21,100
Nitrofurantoin	Nitrofurantoin reduction	400	13,807
Nitrofurazone	Nitrofurazone reduction	400	15,780
Quinone	NADPH reduction	340	6,200

Table 2-3: Spectroscopic details of reactions followed in Steady-State enzyme kinetics experiments.

2.6.1.3 Determination of kinetic parameters

For each reaction the initial rate (v_i) was calculated for a range of concentrations of one substrate (A) whilst keeping the concentration of the other substrate (B) constant. The concentrations used of the variable substrate ranged from ~0.2-10 x K_m, unless solubility meant higher concentrations could not be reached. The approximate K_m was initially estimated by use of a direct linear plot (Cornish-Bowden, 2004). The data could then be fitted to the modified single-substrate Michaelis-Menten equation (Equation 2-1); where [E] is the enzyme concentration, [A] the concentration of the variable substrate, v_i the initial rate, k_{catapp} is the apparent k_{cat} and K_{mApp} the apparent K_m of substrate A at that concentration of substrate B. To calculate the specificity

constant (k_{cat}/K_m) for the variable substrate, data sets were fitted to a modified form of the Michaelis-Menten equation (Equation 2-4).

Equation 2-1

$$\frac{v_i}{[E]} = \frac{k_{catapp}[A]}{K_{mAapp} + [A]}$$

This provides an estimate of k_{catapp} and K_{mAapp}, where k_{catapp} is the apparent maximum number of catalytic cycles the enzyme can catalyse per unit time, and K_{mAapp} is the apparent concentration of substrate A that produces an initial rate of half of the maximum initial rate. For a ping-pong reaction the true values of k_{cat} and K_m are dependent of the concentration of the fixed substrate, B, as described by Equation 2-2 and Equation 2-3 (hence the values at single substrate concentrations are described as k_{catapp} and K_{mAapp}).

Equation 2-2

$$K_{mAapp} = \frac{K_{mA}[B]}{K_{mB} + [B]}$$

Equation 2-3

$$k_{catapp} = \frac{k_{cat}[B]}{K_{mB} + [B]}$$

The second order rate constant (k_{cat}/K_m) (Engel, 1981) is independent of the concentration of the fixed substrate for a ping-pong reaction. Where k_{cat} or K_m cannot be calculated, for example when there is no available data for high substrate

concentrations, k_{cat}/K_m can be estimated accurately by fitting to a modified Michaelis-Menten equation (Equation 2-4).

Equation 2-4

$$\frac{v_i}{[E]} = \frac{k_{catapp}[A]}{\left(\frac{k_{catapp}}{k_{cat}/K_m}\right) + [A]}$$

To get a global fit for the nitroreductase bi-bi enzyme mechanism, kinetic measurements were collected over a range of concentrations of both substrates and fit to Equation 2-5 describing the overall kinetics for a ping-pong reaction.

Equation 2-5

$$\frac{v_i}{[E]} = \frac{k_{cat}[A][B]}{K_{mA}[B] + K_{mB}[A] + [A][B]}$$

As for the one substrate equation, the second-order rate constants can be calculated by fitting to a modified equation (Equation 2-6).

Equation 2-6

$$\frac{v_i}{[E]} = \frac{k_{cat}[A][B]}{\left(\frac{k_{cat}}{k_{cat}/K_{mA}}\right)[B] + \left(\frac{k_{cat}}{k_{cat}/K_{mB}}\right)[A] + [A][B]}$$

2.6.1.4 Inhibition studies

Reactions were recorded as described for each substrate, but each experiment was repeated at various concentrations of the inhibitor. The type of inhibition relative to the first and second substrates was determined by its affect on the apparent Michealis-Menten parameters. The inhibition constants, K_{iU} for uncompetitive inhibition, and K_{iC} for competitive inhibition could then be estimated by fitting all the datasets at each inhibitor concentration to Equation 2-7 and Equation 2-8 respectively. The inhibition constants of both K_{iU} and K_{iC} for mixed inhibition were calculated using Equation 2-9.

Equation 2-7

$$\frac{v_i}{[E]} = \frac{\left(\frac{k_{catapp}}{1 + {}^{[I]}/_{K_{iU}}}\right)[S]}{\left(\frac{K_{mapp}}{1 + {}^{[I]}/_{K_{iU}}}\right) + [S]}$$

Equation 2-8

$$\frac{v_i}{[E]} = \frac{k_{catapp}[S]}{K_{mapp}\left(1 + {}^{[I]}/_{K_{iC}}\right) + [S]}$$

Equation 2-9

$$\frac{v_{\mathrm{i}}}{[\mathrm{E}]} = \frac{k_{\mathrm{catapp}}[\mathrm{S}]}{K_{\mathrm{mapp}}\left(1 + {}^{[\mathrm{I}]}/_{K_{\mathrm{iC}}}\right) + [\mathrm{S}]\left(1 + {}^{[\mathrm{I}]}/_{\mathrm{K_{iU}}}\right)}$$

2.6.2 Stopped-Flow Enzyme Kinetics

2.6.2.1 Apparatus

Stopped-flow experiments on engineered NfsB mutants were performed on an Applied Photophysics DX17MV spectrophotometer. The machine was set in absorbance mode with an optical pathlength of 1 mm. The monochromator slit widths were set to 4 mm. The two reaction components were split into the two drive syringes of 1 ml each, and the concentrations were halved upon mixing. A stop volume of 120 µl was used.

NfsA stopped-flow experiments were performed at the University of Bristol, UK; on an Applied Photophysics SX20 stopped flow spectrophotometer, with the help of Dr Ross Anderson and Dr Craig Armstrong. Experiments were performed within an anaerobic Belle technology glove box. A stop volume of 200 µl was used. A single wavelength spectrophotometer mode was used for monitoring the reduction of the flavin by NADPH with monochromater slit widths of 0.5 mm. For the reductive half reactions, a photodiode array was used and monochromater slit widths were set to 2 mm.

2.6.2.2 Test reaction and determination of instrument Deadtime

The mixing chamber and observation chamber of the stopped-flow instrument are separated within the spectrophotometer setup, thus it is inevitable that some part of the reaction will take place in the time it takes to travel between them. So for accurate readings to be taken this 'deadtime' must be accurately known so that data before this point is not used in data fitting and the amount of product formed by the reaction can be determined by extrapolation.

A suitable test reaction has been described involving the simple first-order reaction of L-ascorbate with 2,6-dichlorophenolindophenol (DCIP), performed at pH 2.0 to minimise errors (Tonomura *et al.*, 1978) (Figure 2-4). All reactions had final concentrations of 250 µM DCIP, 200 mM NaCl, 10 mM HCl and 5% propanol. The final concentration of L-ascorbic acid ranged from 1 – 60 mM. L-ascorbate was in a large excess over DCIP to ensure true first-order kinetics. The reaction was monitored by monitoring the disappearance of DCIP at 524 nm. Five reactions were run at each L-ascorbate concentration and the results were averaged, the results are shown in Figure 2.5. Data sets for each run were fitted to Equation 2-10 with fitting limited to data post 2 ms (to ensure no bias from data during the dead time), where A is the absorbance (A_0 the initial absorbance after mixing), C the amplitude, k_{obs} the observed rate constant, and t time.

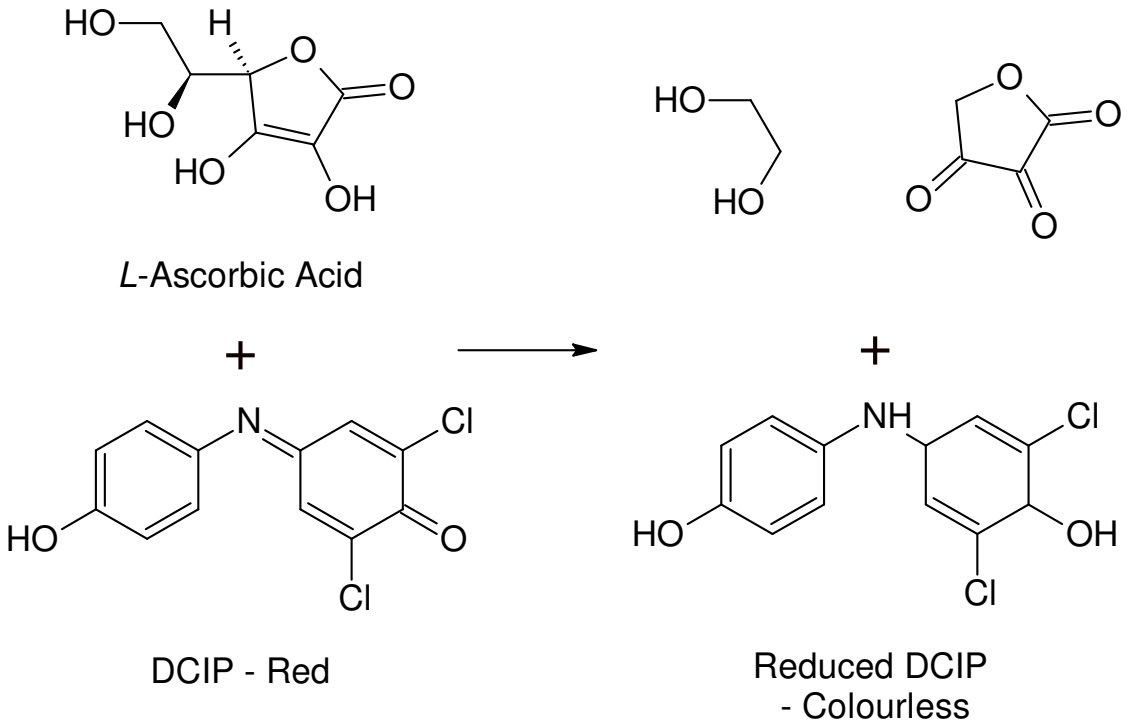


Figure 2-4: Reaction of DCIP and *L*-ascorbic acid. The reaction monitored during the test reaction of the stopped-flow apparatus.

Equation 2-10

$$A = A_0 + Ce^{-k_{obs}t}$$

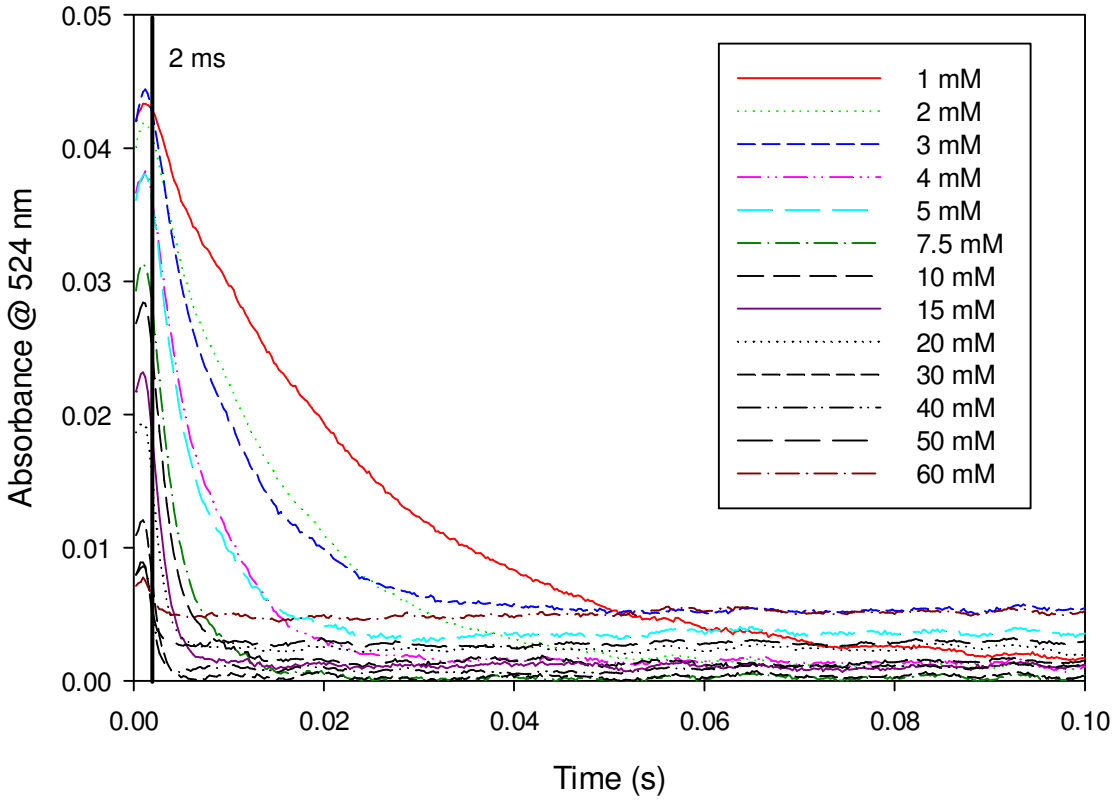


Figure 2-5: The reduction of DCIP (monitored by absorbance change at 524 nm) by various L-ascorbic acid concentrations. A black line is shown indicating the 2 ms point. Only data to the right of this line was used for fitting.

If the calculated rate constants are plotted against the concentration of ascorbate (Figure 2-6) then a straight line is seen which begins to curve off at high concentrations of ascorbate (this can be seen in Figure 2-6 as where the standard error increases, the value for 60 mM ascrobate is missing as the reaction was complete within the deadtime). This plot gives an estimation of the maximum rate that the instrument can observe (~1500 s^{-1}).

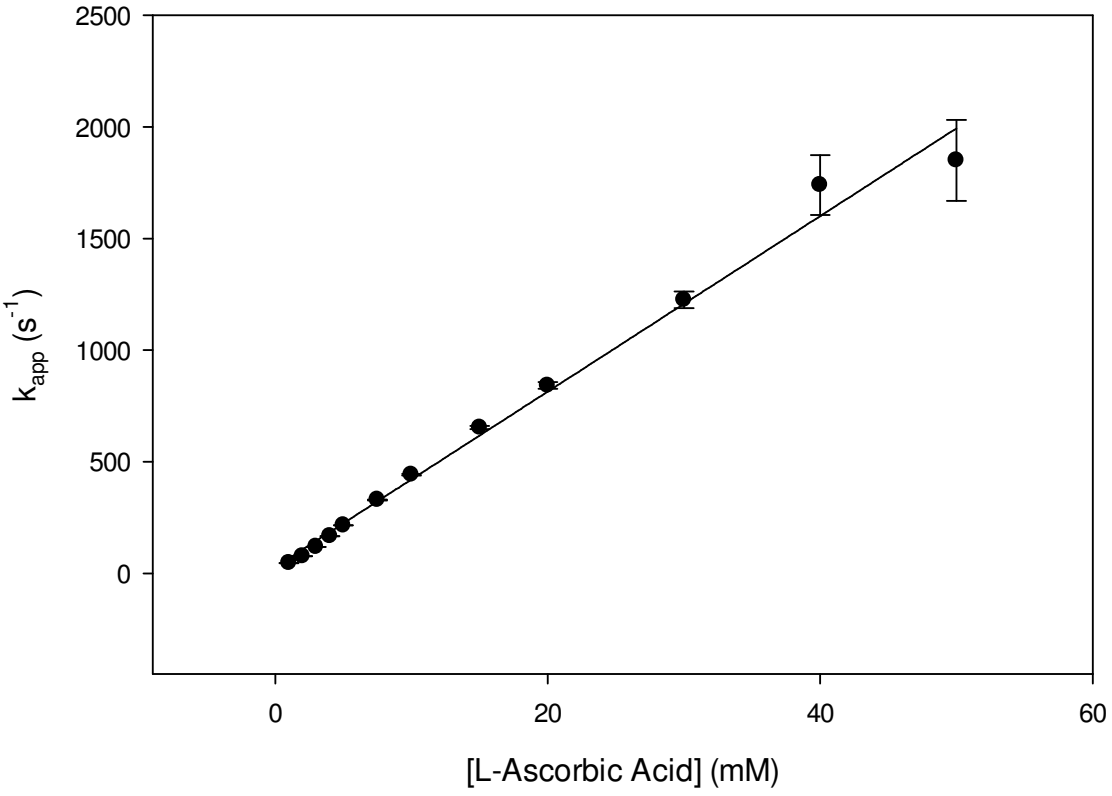


Figure 2-6: Effect of increasing ascorbate concentration on the apparent rate constant (k_{app}) of the reaction with DCIP monitored at 524 nm.

The dead time can be defined by Equation 2-11, where t_d is the dead time. From this we can get Equation 2-12, where ΔA_{obs} is the observed amplitudes and ΔA_{tot} is the actual total absorbance change of the reaction. This equation allows a linear plot of $\ln(\Delta A_{obs})$ against k_{app}, where the gradient of the line gives the deadtime and the intercept with the y-axis gives the natural log of the absorbance at the start point of the reaction. The plot for the data shown above is shown in Figure 2-7. The gradient of the graph gives the deadtime of the instrument at 1.0 ms. The instrument at Bristol also had a deadtime of 1.0 ms (Dr Ross Anderson, personal communication). Therefore for all stopped-flow kinetics studies, data points before this were not included in fitting.

Equation 2-11

$$\frac{t_d}{t_{\frac{1}{2}}} = \ln\left(\frac{x_{tot}}{x_{obs}}\right)/ln2$$

Equation 2-12

$$\ln(\Delta A_{\text{obs}}) = \ln(\Delta A_{tot}) - k_{app}.t_d$$

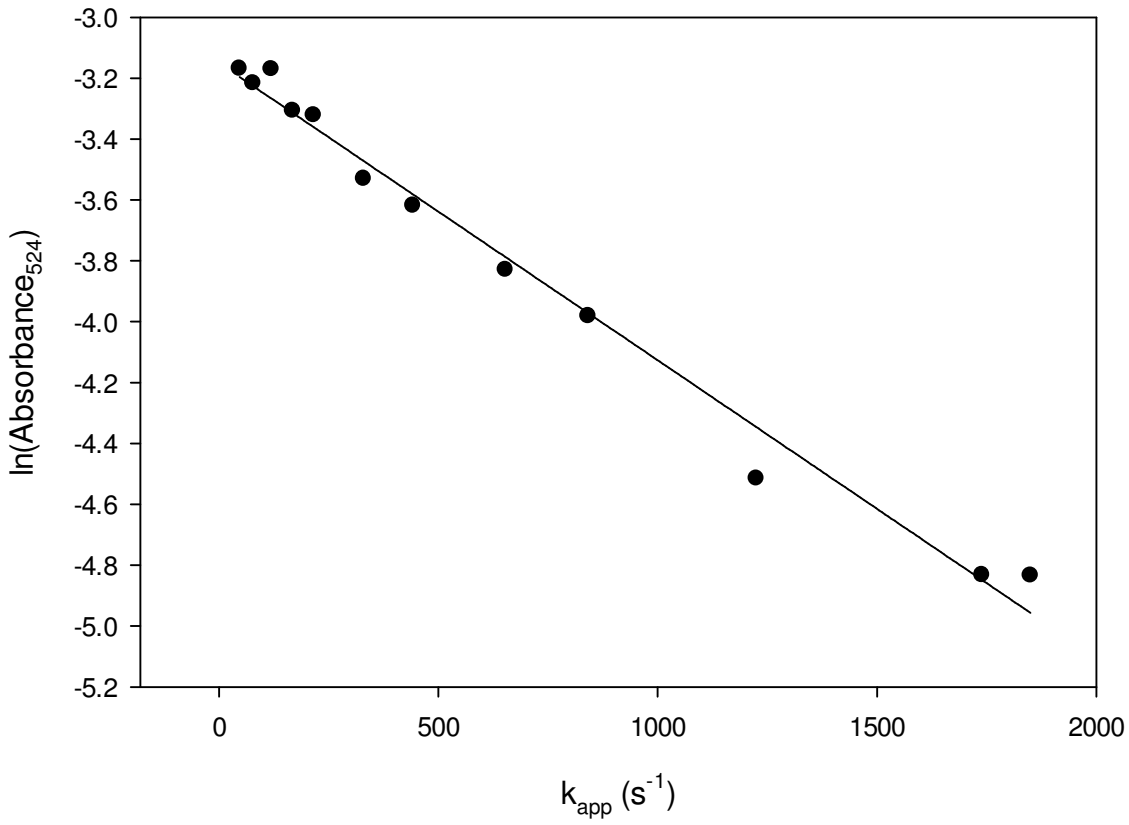


Figure 2-7: Graph used for deadtime calculation of the stopped-flow instrument at the University of Birmingham.

2.6.2.3 Experimental conditions for stopped-flow experiments

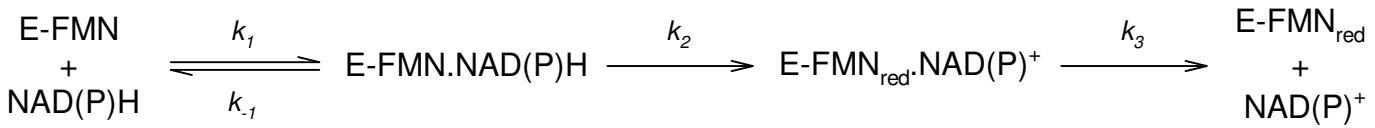
Experiments at Bristol were performed at 25°C, and at Birmingham at room temperature (~24-25°C), all in 10 mM Tris pH 7.0. In all experiments, one syringe contained enzyme and the other substrate. The reaction was followed by monitoring the oxidation/reduction of the enzyme-bound FMN prosthetic group at 454 nm. All enzymes were dialysed into the Tris buffer to remove any excess FMN. Each enzyme was diluted to 10 μM (for a 5 μM final concentration). Keeping the substrate in excess ensured the reaction was pseudo-first-order. To measure the first/oxidative half reaction NAD(P)H concentrations were varied from 25-1000 μM. At each substrate concentration 4-8 duplicate reactions were performed. For the second/reductive half reaction the enzyme FMN complex was reduced by slow addition of dithionite within the anaerobic chamber, monitored at 454 nm using an Ocean Optics USB2000+UV-Vis spectrometer, until the solution was almost colourless. All substrates were made up in buffer containing 10% DMSO (5% final). Final concentrations of CB1954 ranged from 25-1000 μM, and menadione from 10-500 μM.

2.6.2.4 Determination of kinetic parameters

Figure 2-8 shows the breakdown of the reaction scheme of the bi-bi ping-pong reaction catalysed by both the *E. coli* nitroreductase enzymes. Each individual step can be seen and the rate constants are labelled. Each half reaction has a dissociation constant (K_{dA} for the first substrate, NAD(P)H, and K_{dB} for the second substrate, e.g. CB1954) shown in Equation 2-13 for the oxidative half reaction and Equation 2-14 for the reductive half reaction. For the purposes of the stopped-flow experiment the assumption is made that the colour change is caused by the oxidation/reduction step

and not by the release step. So it is assumed the rates we see equate to k_2 (or k_5), and are not a combination of k_2 and k_3 (or k_5 and k_6).

a) - Oxidative Half-Reaction



b) - Reductive Half-Reaction

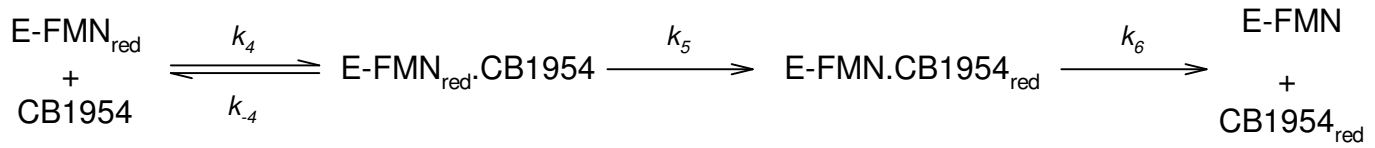


Figure 2-8: Reaction scheme of the bi-bi ping-pong reaction catalysed by the *E. coli* nitroreductases using the example of the reduction of a nitro group to a nitroso. a) shows the oxidative half reaction and b) the reductive half reaction.

Equation 2-13

$$K_{\mathrm{dA}} = \frac{k_1}{k_{-1}}$$

Equation 2-14

$$K_{\mathrm{dB}} = \frac{k_4}{k_{-4}}$$

2.6.2.4.1 Oxidative half-reaction

For the oxidation of NAD(P)H by E-FMN, the absorbance of FMN was monitored and this decreases as the reaction progresses. Each individual plot of absorbance against time was fit to Equation 2-10, a single exponential, giving a pseudo-first-order rate constant, *k*, for each plot. These rate constants were then plotted against substrate concentration and fitted to a hyperbola, which is analogous to the Michaelis-Menten equation (Equation 2-15), where k_{obs} is the observed rate constant and A is the first substrate. In order to calculate the k_2/K_{dA}, the data was fitted to the modified Equation 2-16.

Equation 2-15

$$k_{\mathrm{obs}} = \frac{k_2[\mathrm{A}]}{K_{\mathrm{dA}} + [\mathrm{A}]}$$

Equation 2-16

$$k_{\mathrm{obs}} = \frac{k_2[\mathrm{A}]}{\left(\frac{k_2}{k_2/K_{\mathrm{dA}}}\right) + [\mathrm{A}]}$$

2.6.2.4.2 Reductive half-reaction

For the reductive half reaction, the absorbance of FMN increases so each reaction was fitted to a single exponential (Equation 2-17). Then as for the oxidative half reaction the rate data could be fitted to hyperbolas, Equation 2-18 and Equation 2-19, where [B] is the concentration of the second substrate.

Equation 2-17

$$A = A_0 + C(1 - \mathrm{e}^{-k_{\mathrm{obs}}t})$$

Equation 2-18

$$k_{\mathrm{obs}} = \frac{k_5[\mathrm{B}]}{K_{\mathrm{dB}} + [\mathrm{B}]}$$

Equation 2-19

$$k_{\mathrm{obs}} = \frac{k_5[\mathrm{B}]}{\left(\frac{k_5}{k_5/K_{\mathrm{dB}}}\right) + [\mathrm{B}]}$$

3. NfsB Mutant Structures

3.1 The Structure of NfsB

The *E. coli* minor nitroreductase, NfsB, has been analysed using X-ray crystallography with the aiming of understanding the chemical environment of the active site to guide future design of prodrugs or mutants improving the efficiency of the combination for GDEPT (Section 1.5). The structure of wild-type NfsB has been published previously both in its native form (Parkinson *et al.*, 2000) and complexed with inhibitors nicotinic acid (Lovering *et al.*, 2001) and acetate and the substrate nitrofurazone (Race *et al.*, 2005) (Section 1.6.2.2). The overall structure of the NfsB dimer is shown in Figure 1-23 and the FMN binding is detailed in Figure 1-24. Structures have also been determined for several single mutants and a double mutant (Race *et al.*, 2007). The most improved single, double and triple mutants for the reduction of CB1954 have been studied by stopped-flow and steady-state kinetics (Jarrom *et al.*, 2009). In this work the structures of the most active double (T41L/N71S), triple (T41Q/N71S/F124T) and quadruple (T41Q/N71S/F124T/M127V) mutants are determined by X-ray crystallography. The structures of each of these mutants, complexed with the inhibitor and NADH analogue, nicotinic acid, are presented here. Also presented are structures of several of these mutants bound to citrate or acetate, and other structures found in previously uncharacterized crystal forms. And finally the results of attempts to co-crystallize the mutant proteins with CB1954 and related prodrugs are discussed.

3.2 Purity of purified samples

The NfsB WT, T41L/N71S, T41Q/N71S/F124T and T41Q/N71S/F124T/M127V proteins were grown and purified as described in Section 2.2, using ammonium sulphate precipitation followed by Phenyl-Sepharose and Q-Sepharose chromatography, as described previously (Lovering *et al.*, 2001). SDS-PAGE gels are shown for the double mutant fractions from the phenyl-sepharose column are shown in Figure 3-1 a), and from the Q-sepharose in Figure 3-1 b).

a) b)

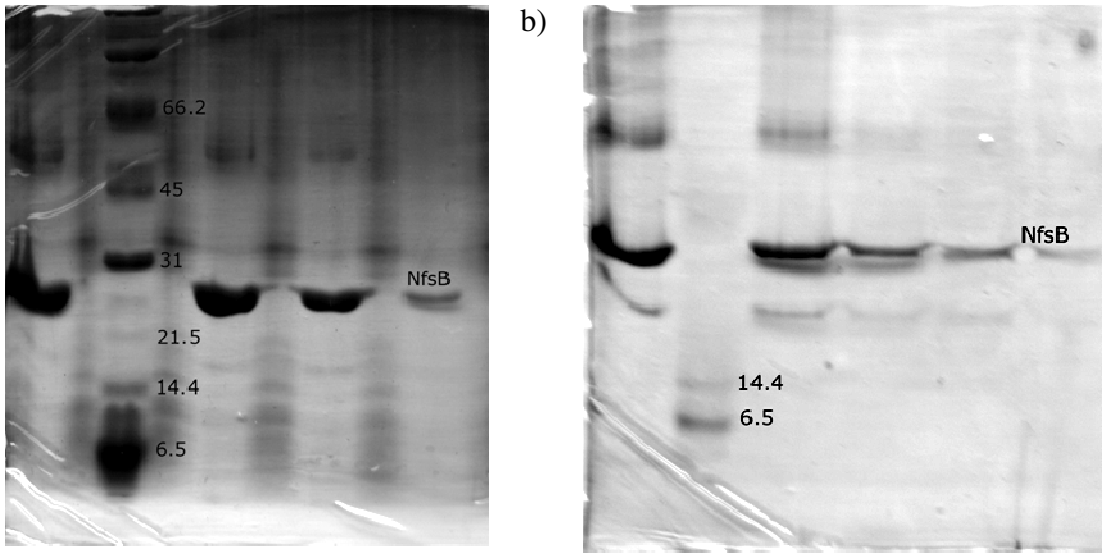


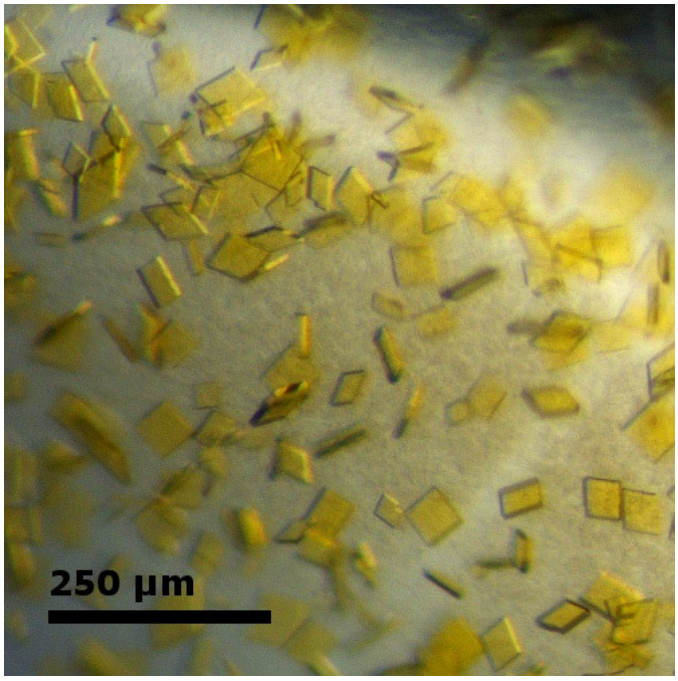
Figure 3-1: SDS-PAGE gels of T41LN71S NfsB mutant fractions a) after purification using a Phenyl Sepharose matrix and b) after subsequent purification using a Q Sepharose matrix. A marker lane is shown (second from the left on each image), and visible bands are labelled for size in kDa. The other lanes contain analysis of various pooled fractions of protein.

3.3 Crystal Growth at pH 4.6

Initial conditions were screened based upon the published NfsB crystallization conditions (Lovering *et al.*, 2001) at pH 4.6 containing PEG 4,000 and sodium acetate.

The first crystals were seen in sitting drop trays and were small yellow plates (Figure 3-2 a)). The hanging drop method produced more reproducible crystal growth and, once optimised for drop size and constituent concentrations, produced much larger crystals (up to 500 x 500 x 40 µm plates) (Figure 3-2 b)). All crystals appeared within 24 hours and reached full size within 72 hours. All crystals in this acetate buffer were tetragonal yellow plates of various thickness, the most common form previously observed (Lovering *et al.*, 2001). There was no sign in any of the > 50 crystal trays set up in these conditions of either the monoclinic or orthorhombic crystal forms previously described. When attempting to co-crystallize with ligands in these conditions, crystals did not grow in the presence of high concentrations of most ligands that were tested. These included NFZ, NFN, menadione, quinone and CB1954, where crystals did not grow in the presence of more than 5 mM. Where crystals did grow in the presence of low ligand concentrations, no ligand density was observed in the active site.

a)



b)

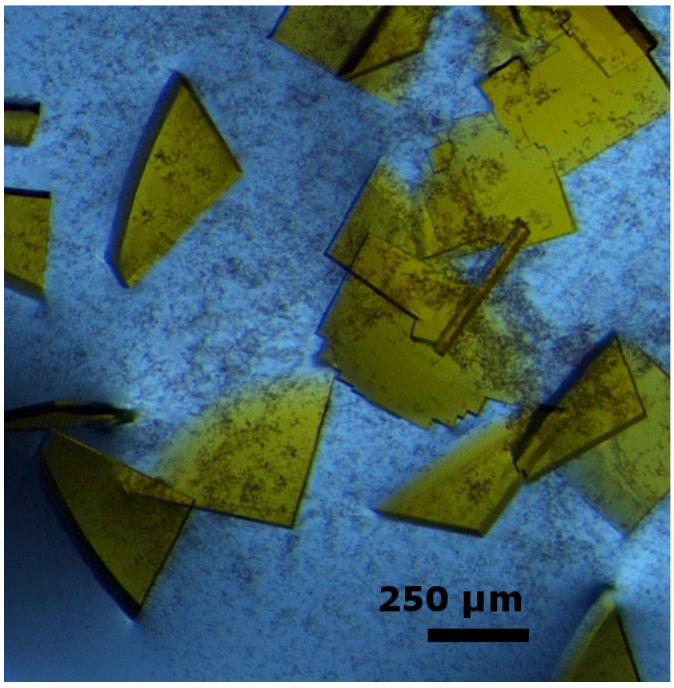


Figure 3-2: Crystals of NfsB mutants grown at pH 4.6. a) Initial crystal growth, T41Q/N71S/F124T mutant crystals grown via sitting drop method. b) Optimised crystal growth, T41Q/N71S/F124T mutant crystals grown via hanging drop method.

Attempts were made to repeat the successful soaking experiments used to solve the structure of WT-NfsB bound to nitrofurazone (Race *et al.*, 2005). In these experiments the acetate buffer was replaced with a citrate buffer (both at pH 4.6), to remove the acetate from the active site. This was not successful with the crystal forms shown here, which are all tetragonal (the previously successful soak was of a monoclinic form), however a crystal frozen after the citrate soak step did produce a structure containing citrate in the active site.

3.4 Crystal Growth at pH 7.0

It is most desirable to obtain crystals of NfsB at neutral pH, as this is close to the conditions in which it would be working in its GDEPT setting. Also when attempting to obtain NfsB/CB1954 complexes, the aziridine ring would rapidly ring open at the lower pH (pH 4.6), and so any complex would not be physiologically relevant. Initial screens at neutral pH were based around the published conditions of Parkinson *et al.* (Parkinson *et al.*, 2000), containing PEG 8,000 and potassium phosphate. Crystals grown in these conditions initially grew in two forms; plates as with the acetate conditions (although these were much smaller and thinner), and needles (Figure 3-3 a)).

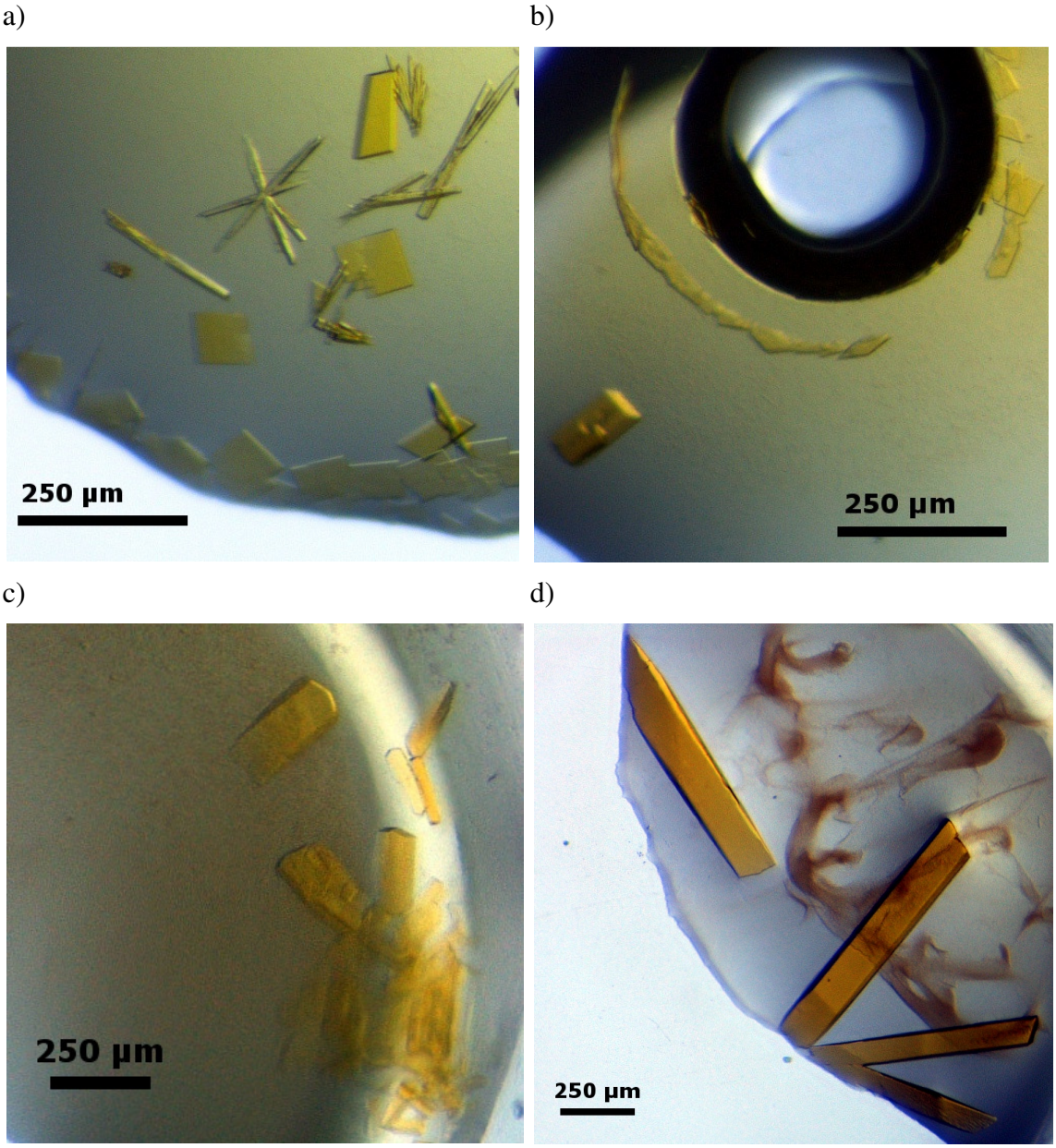
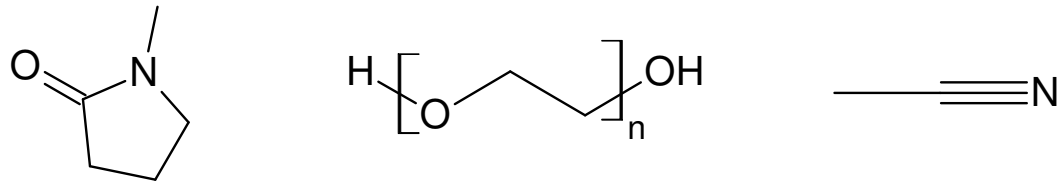


Figure 3-3: Crystals of NfsB mutants grown at pH 7.0. a) Initial crystals showing multiple crystal forms of T41L/N71S. b) T41L/N71S mutant crystals co-crystallized with SN27683, tetragonal plates can be seen as well as a larger monoclinic form (also seen in d)). c) T41L/N71S double mutant crystals co-crystallized with SN27686, a thin monoclinic form. d) T41Q/N71S/F124T mutant crystals co-crystallized with SN27686, large wedge-like monoclinic crystals.

The first crystals grown at pH 7.0 were fragile and so the decision was made to co-crystallize with the cryo-protectant ethylene glycol (at 15-20 %) to remove the necessity of handling the crystals before mounting. These crystal of NfsB were also temperature sensitive as seen when they dissolved after a failure in the crystal room allowed the temperature to rise from 18°C to 22°C. Attempts to crystallize at room temperature were unsuccessful and at 4°C the protein precipitated too readily.

In an attempt to obtain protein complexes with CB1954 and related prodrugs, co-crystallization experiments were undertaken. Previous attempts in this lab with WT-NfsB had failed to give a complex, but it should be possible with the mutants as they bind CB1954 much more tightly than WT (Section 1.6.2.3). Crystals grew in the presence of up to 10 mM of each prodrug (see Table 3-1 for an idea of the solvent concentrations used). In the presence of CB1954 (in a 2:7 NMP:PEG300 solvent - Figure 3-4), tetragonal plates were seen, as before. However in the presence of the other dinitrobenzamide prodrugs, three new crystal forms were also observed (Figure 3-3 b-d)) (in acetonitrile (Figure 3-4), see Table 3-1 for an idea of the concentrations). The most common of these new forms was a monoclinic form which resembled a thick wedge (Figure 3-3 b) and d)), observed primarily in the presence of SN27683 and SN27686. These crystals were robust, easy to handle and diffracted to around 1 Å, greater than any published data for NfsB. Another monoclinic space group, which formed smaller crystals, was seen with most of the dinitrobenzamide prodrugs but was much more fragile (Figure 3-3 c)). These crystals diffracted to around 1.6 Å, similar to the maximum resolutions previously observed (Lovering *et al.*, 2001). Finally a very

thin uneven crystal form, which was only a pale yellow, was seen only with the T41LN71S mutant in the presence of SN29222 (not shown). This plate-like form did diffract but the crystals were highly mosaic which made indexing impossible.



N-methylpyrrolidone (NMP) | Polyethylene glycol (PEG) | Acetonitrile

Figure 3-4: The chemical structures of the solvents used for CB1954 and the other dinitrobenzamide prodrugs. PEG 300 is a combination of polyethylene glycol molecules with an average molecular weight of 300 Da.

3.5 Summary of datasets

Datasets were collected for three mutants of NfsB: the most active double mutant to date (T41L/N71S), the most active triple mutant (T41Q/N71S/F124T) and the most active quadruple mutant (T41Q/N71S/F124T/M127V) with different ligands. Nearly 100 datasets were collected and 12 were fully refined and analysed. Structures for all three mutants were solved complexed to the inhibitor nicotinic acid. The structure of the triple mutant was also determined with bound acetate and with the tricarboxylic acid citrate, as was that of the double mutant with nothing bound in the active site. All these were from crystals grown at pH 4.6. The results of co-crystallizing the mutants

with CB1954 and other dinitrobenzamide prodrugs and neutral pH are shown. The crystallization conditions are summarised in Table 3-1, the data collection statistics in Table 3-2, the processing statistics in Table 3-3 and the refinement statistics in Table 3-4. The datasets shown here include previously unobserved crystal forms of NfsB, including one that provided a structure at higher resolution than any previously seen, a primitive monoclinic form in the space group P 1 2_1 1 seen with the triple mutant when co-crystallized with SN27683 and SN27686. The highest resolution NfsB structure in the pdb to date is at 1.7 Å (Lovering *et al.*, 2001, Race *et al.*, 2005), whilst the highest resolution structure shown here is at 1.08 Å.

	Crystallization						Soak	
Dataset Identifier	*Tray Type*	*Precipitant*	*Buffer*	*Additive*	*Cryo. Prot.*	*Protein Conc. (mg/ml)*	*Cryo. Prot.*	*Ligand*
Double Nicotinate	Sitting Drop	18% PEG 4,000	200 mM Na Acetate	15 mM Nicotinic Acid	-	12	20% Ethylene Glycol	-
Triple Nicotinate	Hanging Drop	14% PEG 4,000	100 mM Na Acetate	20 mM Nicotinic Acid	15% Ethylene Glycol	10	-	-
Quad Nicotinate	Sitting Drop	12% PEG 4,000	100 mM Na Acetate	15 mM Nicotinic Acid	-	6	15% Ethylene Glycol	-
Double Empty	Hanging Drop	13% PEG 4,000	100 mM Na Acetate	15 mM Nicotinic Acid	15% Ethylene Glycol	11	20% Ethylene Glycol	100 mM Citrate
Triple Acetate	Hanging Drop	13% PEG 4,000	100 mM Na Acetate	15 mM Nicotinate Acid	15% Ethylene Glycol	13	20% Ethylene Glycol	25 mM Citrate
Triple Citrate	Sitting Drop	13% PEG 4,000	100 mM Na Acetate	15 mM Nicotinic Acid	15% Ethylene Glycol	10	15% Ethylene Glycol	100 mM Citrate
Double CB1954	Sitting Drop	30% PEG 8,000	10 mM K Phosphate	8.5 mM CB1954	15% Ethylene Glycol	15	2.7% NMP*	9.3% PEG 300*
Double SN27419	Sitting Drop	23% PEG 8,000	10 mM K Phosphate	8 mM SN27419	15% Ethylene Glycol	11	36% Acetonitrile*	-
Double SN27683	Sitting Drop	28% PEG 8,000	20 mM K Phosphate	12 mM SN27683	15% Ethylene Glycol	13	20% Acetonitrile*	-
Triple SN27683	Sitting Drop	25% PEG 8,000	10 mM K Phosphate	8 mM SN27683	15% Ethylene Glycol	14	20% Acetonitrile*	-
Triple SN27686	Sitting Drop	28% PEG 8,000	20 mM K Phosphate	20 mM SN27686	15% Ethylene Glycol	11	36% Acetonitrile*	-

Table 3-1: Crystallization conditions for crystals giving rise to NfsB mutant datasets. All acetate buffers are at pH 4.6, and all phosphate buffers at pH 7.0. * Further contents of the mother liquor used to solubilise prodrugs rather than soak contents. These crystals were mounted straight from the drop.

	General					Unit Cell				
Dataset Identifier	*Active Site Contents*	*X-ray Source*		*Wavelength (Å)*	*Degrees of Data (°)*	*Space Group*	*a (Å)*	*b (Å)*	*c (Å)*	*β (°)*
Double Nicotinate	Nicotinate	ESRF	ID14-1	0.934	90	P 4_1 2_1 2	56.9	56.9	264	-
Triple Nicotinate	Nicotinate	ESRF	ID23-1	1.00	80	P 4_1 2_1 2	57.6	57.6	265	-
Quad Nicotinate	Nicotinate	ESRF	ID14-1	0.934	120	P 4_1 2_1 2	57.3	57.3	266	-
Double Empty	None	ESRF	ID14-1	0.934	90	P 4_1 2_1 2	57.2	57.2	263	-
Triple Acetate	Acetate	ESRF	ID14-1	0.934	81	P 4_1 2_1 2	57.3	57.3	263	-
Triple Citrate	Citrate	ESRF	ID14-1	0.934	90	P 4_1 2_1 2	57.3	57.3	264	-
Double CB1954	Unidentified	Diamond	I03	0.976	83	P 4_1 2_1 2	56.9	56.9	257	-
Double SN27419	Unidentified	Diamond	I03	0.920	139	P 4_1 2_1 2	57.5	57.5	267	-
Double SN27683	Unidentified	Diamond	I03	0.920	375	C 1 2 1	115	43.0	95.0	115
Triple SN27683	Unidentified	Diamond	I03	0.976	153	P 1 2_1 1	54.1	60.6	75.0	99.9
Triple SN27686	Unidentified	Diamond	I04	0.979	152	P 1 2_1 1	54.0	50.6	75.0	100

Table 3-2: Data collection statistics for NfsB mutant datasets. ESRF refers to the European Synchrotron Radiation Facility, Grenoble, France, whilst Diamond refers to Diamond Light Source, UK.

	Processing							
Dataset Identifier	*Resolution Range (Å)*	*No. Observations*	*No. Unique*	*Multiplicity (%)*	*Completeness (%)*	*I/if*	R_{sym} *(%)*	*Mosaicity*
Double Nicotinate	47.71-2.20 (2.33-2.20)	172,983 (10,839)	15,943 (1,068)	11 (10)	98.4 (92.5)	24.9 (5.2)	7.5 (53.2)	0.214
Triple Nicotinate	48.24-1.65 (1.74-1.65)	290,689 (18,810)	53,617 (6,749)	5.4 (2.8)	97.7 (86.8)	13.2 (1.9)	7.8 (55.7)	0.124
Quad Nicotinate	35.04-1.99 (2.10-1.99)	280,907 (19,436)	31,129 (2,176)	9.0 (8.9)	98.9 (95.4)	25.3 (6.2)	7.5 (55.0)	0.141
Double Empty	47.92-1.69 (1.80-1.69)	647,523 (11,240)	49,521 (3,178)	13 (3.5)	99.0 (87.5)	36.9 (3.4)	5.0 (34.8)	0.096
Triple Acetate	47.85-1.60 (1.69-1.60)	345,900 (37,708)	57,801 (8,552)	5.9 (4.4)	98.6 (92.3)	27.2 (5.1)	4.5 (29.3)	0.112
Triple Citrate	47.99-2.09 (2.21-2.09)	189,497 (30,010)	26,461 (4,003)	7.2 (7.5)	99.8 (98.2)	21.6 (5.5)	8.3 (36.9)	0.105
Double CB1954	29.36-1.60 (1.69-1.60)	305,672 (18,289)	54,734 (6,308)	5.6 (2.9)	95.9 (78.8)	18.7 (4.2)	5.9 (22.4)	0.104
Double SN27419	29.73-1.90 (2.00-1.90)	409,427 (56,821)	41,786 (5,899)	9.8 (9.6)	99.7 (98.4)	19.2 (1.6)	7.0 (66.2)	0.155
Double SN27683	28.80-1.54 (1.63-1.54)	294,536 (37,205)	61,162 (8,362)	4.8 (4.4)	97.1 (91.7)	14.0 (2.9)	4.8 (35.1)	0.205
Triple SN27683	27.74-1.08 (1.14-1.08)	472,510 (40,453)	155,634 (17,812)	3.0 (2.3)	91.5 (72.1)	9.3 (1.6)	6.6 (78.1)	0.107
Triple SN27686	27.73-1.10 (1.16-1.10)	488,421 (54,891)	146,872 (17,480)	3.3 (3.1)	91.0 (74.5)	13.1 (3.1)	6.4 (50.1)	0.197

Table 3-3: Processing statistics of NfsB mutant datasets. Values in parentheses are for the highest resolution shell. $R_{sym} = (\Sigma_{hkl}\Sigma_j|I_{hkl,j} - \langle I_{hkl}\rangle|)/\Sigma_{hkl}\Sigma_j I_{hkl,j}$.

	Refinement						Mean B Factors (Å^2)			
Dataset Identifier	*No. of Reflections*	*No. Water Molecules*	*R factor (%)*	*R_{free} (%)*	*RMSD bond angles (°)*	*RMSD Bond Lengths (Å)*	*Protein*	*Ligand*	*FMN*	*Water*
Double Nicotinate	22,344	58	18.31	26.29	0.81	0.005	50.7	60.1	38.7	50.1
Triple Nicotinate	53,495	582	17.84	22.69	0.93	0.006	18.9	21.2	13.2	33.6
Quad Nicotinate	31,128	190	16.18	21.32	0.93	0.007	36.6	53.7	22.3	42.0
Double Empty	49,533	534	15.46	18.68	0.91	0.006	17.4	-	10.9	31.9
Triple Acetate	57,878	497	15.42	18.85	0.91	0.006	16.8	-	10.6	29.3
Triple Citrate	27,018	352	16.4	22.37	0.92	0.007	23.9	84.2	16.6	31.1
Double CB1954	54,623	400	17.53	21.43	0.91	0.006	18.9	17.0	11.0	33.2
Double SN27419	41,671	431	19.25	24.55	0.93	0.006	30.5	44.1	23.6	38.4
Double SN27683	61,115	511	19.46	23.76	0.97	0.006	29.8	31.8	19.7	39.0
Triple SN27683	155,379	567	15.32	16.87	1.22	0.009	18.9	13.7	8.85	28.9
Triple SN27686	140,746	647	12.39	14.33	1.19	0.008	12.8	11.3	7.03	29.3

Table 3-4: Refinement statistics for NfsB mutant datasets. *R-factor* = $(\Sigma| |F_{obs}|-|F_{calc}||)/\Sigma|F_{obs}|$.The *R-free* is the R-factors calculated from the 5% of data left out of refinement.

The structures were solved by molecular replacement using a modified version of the pdb I.D. 1ICR structure (Lovering *et al.*, 2001), which was in the tetragonal crystal form (See Section 2.5.4.2). Most of the datasets have refined well, with R-factors appropriate for the resolution. Statistics were compared to those in the protein data bank using POLYGON (Urzhumtseva *et al.*, 2009). All of the structures presented here compare favourably, although the double nicotinate structure is worse than most. The double nicotinate structure is at the lowest resolution (2.2 Å), and correspondingly has the poorest statistics (highest R-factors and B-factors). The electron density map is less clear than the other structures, but still allowed unambiguous assignment of side chain rotamers and the active site nicotinate. The statistics of the double nicotinate structure are worse due to the high mosaicity and anisotropy of the data. The citrate ligand in the triple citrate structure also has high B-factors relative to the protein, representative of the fact that the density suggests movement and perhaps multiple conformations.

3.6 Overall Structure of NfsB Mutants

The overall structure of the NfsB dimer is shown in Figure 1-23 and discussed in Section 1.6.2.2.1. The monomers of all published structures and all structures presented in this chapter were compared using THESEUS (Theobald and Wuttke, 2008). There are no major conformational differences when the backbones are aligned, but there is some variation (at most 5-10 Å) in the positions of three loops. The most variation is seen with the loop between helices C and D (residues 64-71), which contains the flexible Phe70 residue. Other differences are seen in the loop between helices G and H (residues 129-133) and the loop between sheet 4 and helix G (residues

195-202). The variations were analysed to compare mutants to WT and to contrast different space groups and no patterns were observed, implying that these are all areas of natural variation in the protein structure.

3.7 Overview of Mutant Structures

3.7.1 Mutation Sites

The mutations at Thr41, Asn71 and Phe124 all showed very clear electron density in all 11 structures, allowing unambiguous assignment of side chain rotamers (Figure 3-5 a)-c)). This was true with all the mutant structures presented here. The M127V mutation was less clear (Figure 3-5 d)), although with such a small side chain the possible orientations are fewer.

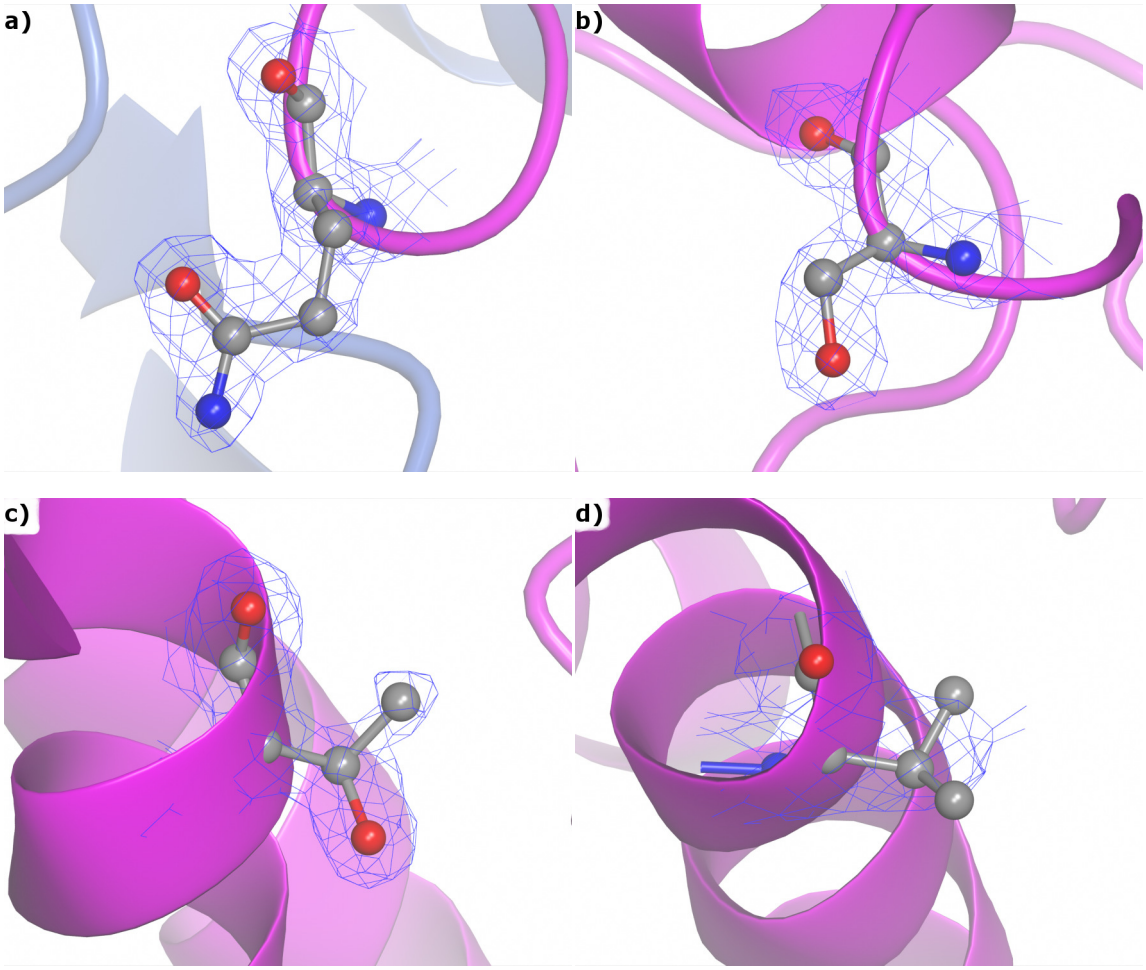


Figure 3-5: Density of mutated residue side chains. a) T41Q mutation, b) N71S, and c) F124T all from the Triple Nicotinate structure, and d) M127V from the Quad Nicotinate structure. Mutated residues are shown in grey, and the blue σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

3.7.2 Summary of Side Chain Differences of Individual Mutations

3.7.2.1 T41L Mutants

In the crystal structure of the T41L single mutant (Race *et al.*, 2007), the conversion of the threonine to the larger leucine caused a slight shift in the surrounding residues (Figure 3-6). This shift is seen in all the double mutant structures presented here, to a

similar degree to those previous structures of the T41L single mutant. The shift is only significant (above the coordinate error of 0.1 Å) for Arg121, Lys122, Phe123 and Phe124, with the maximum change seen in the position of the side chain of Phe124 of just below 1.0 Å. This shift is caused by the effective movement of the side chain of Leu41 upon mutation from threonine, and is a steric effect. When all the monomers in each of the T41L-containing structures are overlaid, the Leu41 residues are well aligned as are most other active site residues. The only residue which shows any difference between all mutant structures containing T41L is Phe123, which rotates by up to 60° about the C^{β} - C1 bond.

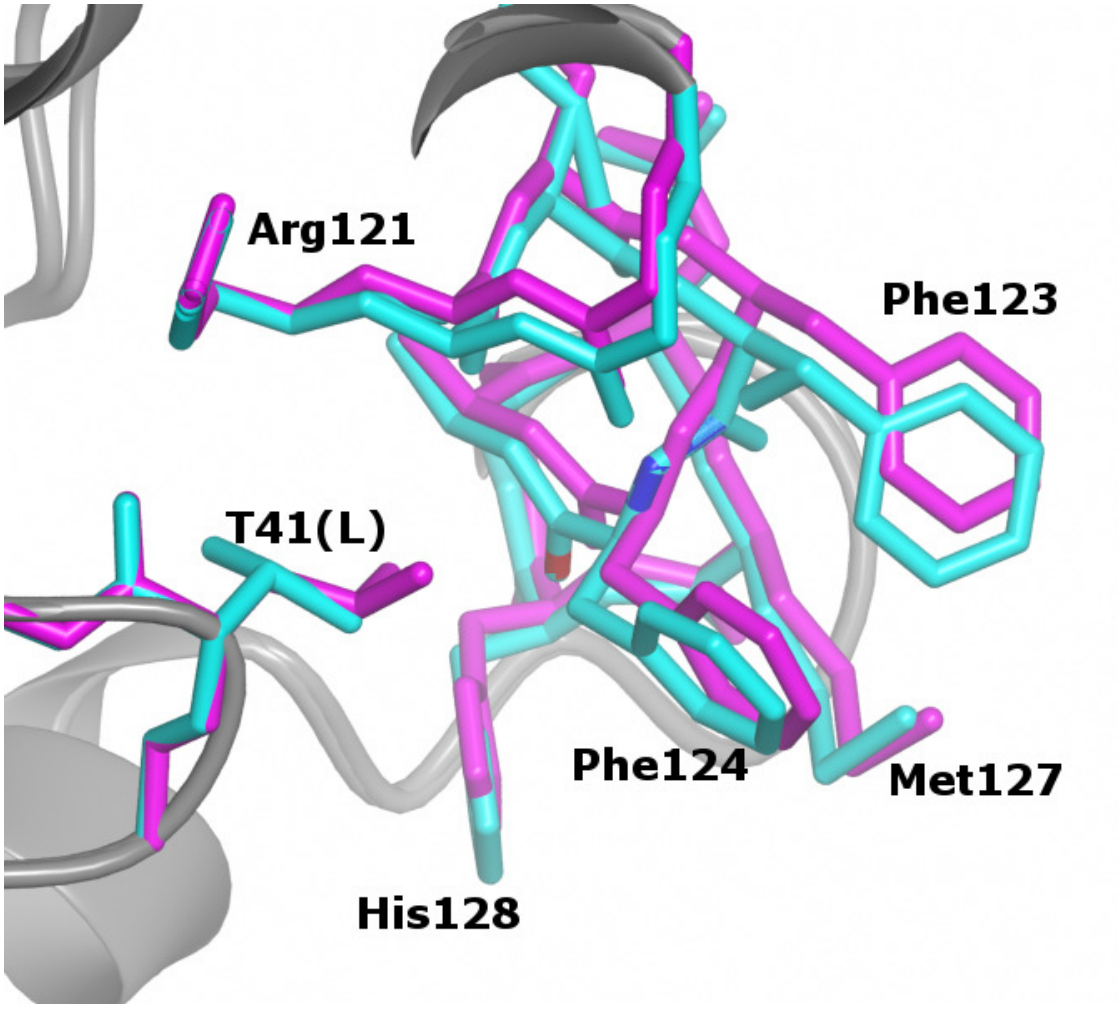


Figure 3-6: Shift of Phe124 and surrounded residues in the presence of the T41L mutation. Comparison of nicotinate-bound WT NfsB (blue – pdb I.D. 1ICR (Lovering *et al.*, 2001)) and T41L/N71S double mutant (purple). Residues within 5 Å of Phe124 are shown. Structures were aligned using THESEUS (Theobald and Wuttke, 2008) and figure was produced using CCP4 Molecular Graphics.

3.7.2.2 T41Q Mutation

All triple and quadruple mutant structures presented in this chapter contain the T41Q mutation. Aligning the monomers of all these structures on the backbone showed that the position of the T41Q was identical in all these structures. Similarly to the T41L,

when the T41Q is compared to the wild-type structure the backbones are perfectly aligned and the only difference is in the side chains, with the smaller threonine replaced with the larger glutamine. There appears to be no other effect on the local environment, other than the region around the 124 residue, as with the T41L. This will be discussed in more detail later.

3.7.2.3 N71S Mutation

Previous crystallographic analysis of the N71S single mutant and the N71S/F124K double mutants (Race *et al.*, 2007) showed that the only changes seen when compared to wild-type were in the Ser71 and neighbouring Phe70. When all the monomers from all N71S-containing mutant structures (i.e. the N71S and N71S/F124K mutants determined previously and the mutants described here) are overlaid, the Ser71 positions are identical in each one, the additional mutations in the double, triple and quaternary mutants having no effect on the Ser71 residue. It was previously suggested that the Phe70 side chain rotamer was changed in the presence of the N71S mutation (Race et al., 2007), however when all these structures are aligned it is seen that the Phe70 side chain can adopt several different rotamers. As was previously observed in the N71S and N71S/F124K mutants, the majority of the structures have the OH hydroxyl group of Ser41 hydrogen bonded to a water molecule, which in turn hydrogen bonds to the Flavin ring and to Lys74 (Figure 1-27). The Ser41 also hydrogen bonds to the backbone carbonyl oxygen of Gly166. This effectively replaces the hydrogen bonding network seen with Asn71 as described previously (Race *et al.*, 2007).

3.7.2.4 F124T Mutation

All F124T mutation sites were clear in each triple mutant structure. The effect of the mutation becomes clear when combined with the T41Q mutation and is discussed in Section 3.9.1.

3.7.2.5 M127V Mutation

The one structure of the quadruple mutant contains two monomers. These overlay well (RMSD = 0.71 Å) and the valines are both in the same conformation. The quadruple mutant will be discussed in more detail later (Section 3.8.4).

3.8 NfsB Double and Triple and Quadruple mutant structures bound to Nicotinate

3.8.1 Structures compared to wild-type

Since all the mutants and wild-type structures have been determined with bound nicotinate this is a good form to compare the effects of the mutations. When all the structures determined here are aligned with all previously determined wild-type NfsB structures it can be seen that, as with previously described mutations (Race *et al.*, 2007), there is little change in the backbone structure of the protein. Small differences are seen in the loop between helices F and G (for labelling see Figure 1-23), which is solvent exposed and the maximum differences between backbone atoms is less than 2.5 Å.

The main differences between the mutants are all seen in the active site. The FMN and nicotinate molecules are all well aligned including the butterfly angle of the flavin rings. It is the side chains of the mutated residues themselves that show the largest differences between the structures (Figure 3-7). The two separate mutations in the T41L/N71S mutant have been seen in previously determined single (T41L and N71S) and double (N71S/F124K) mutant crystal structures (Race *et al.*, 2007), although not with each other.

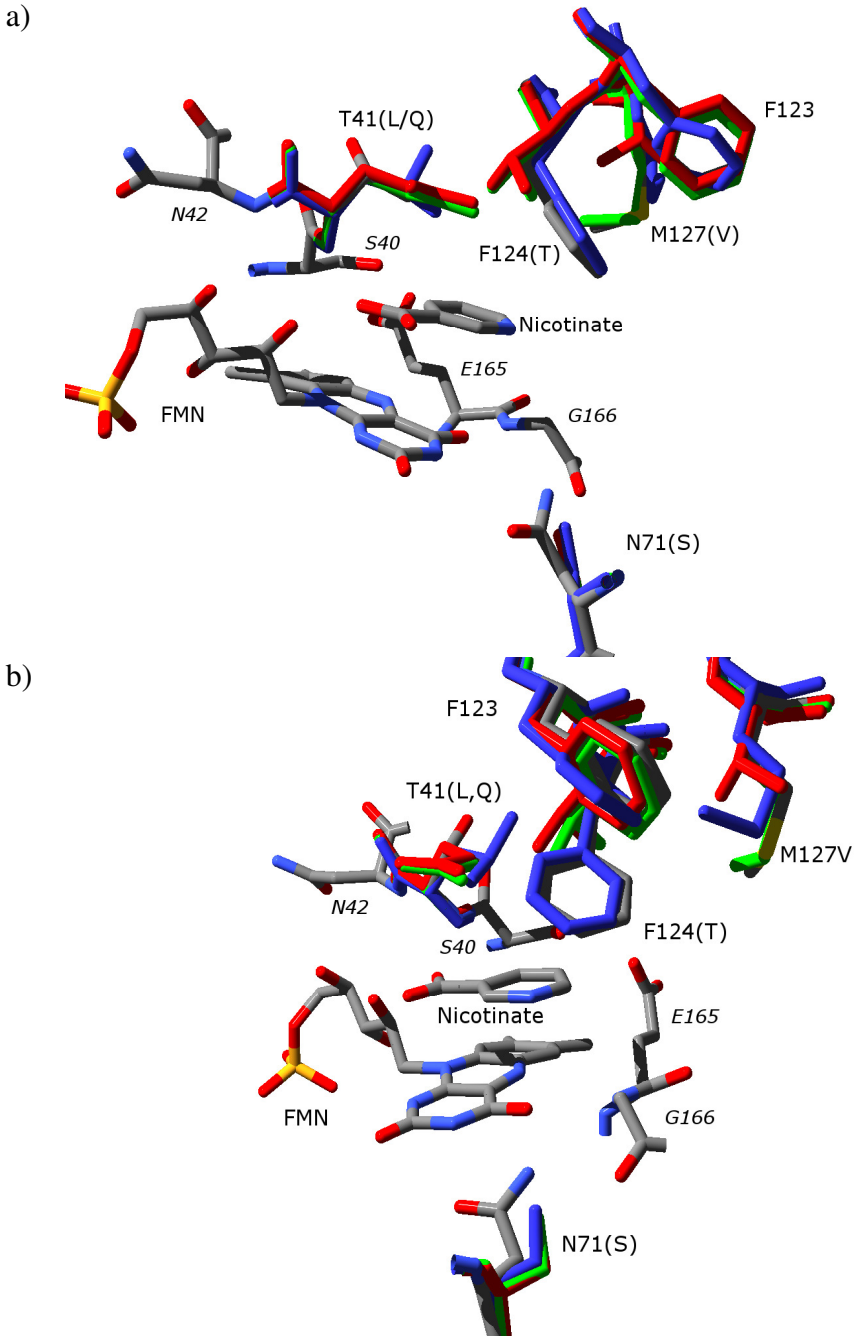


Figure 3-7: Active Site of NfsB, showing a comparison of the mutated residues and other key residues. The wild-type structure (pdb I.D. 1ICR)(Lovering *et al.*, 2001), is coloured according to atom type, whilst only the four residues mutated in this study and Phe123 (as it shows variation across) are shown for three mutants. The double mutant is shown in blue, the triple mutant in green and the quadruple mutant in red. Images a) and b) are 90° apart. Figure produced using Swiss-PdbViewer and POV-Ray.

3.8.2 T41L/N71S Double Mutant Structure

The only differences seen between the nicotinate-bound structure of the wild-type protein and that of the double mutant are those observed for the single mutants (Race *et al.*, 2007), discussed earlier (Section 3.7.2).

3.8.3 T41Q/N71S/F124T Triple Mutant Structure

When comparing the nicotinate-bound wild-type and triple mutant structures, the two align very well indeed (RMSD (all atoms): 0.15 Å – compared to 0.33 Å for both the double and quadruple mutants). In comparison to the double mutant this may be in part due to the higher resolution of the triple mutant structure (1.65 Å compared to 2.20 Å), leading to less ambiguous assignment of side chain positions. The N71S mutation appears to have the same effect as in the double mutant and all previous N71S-containing mutants (Race *et al.*, 2007), as stated previously. However it is the locus of the T41Q and F124T mutations that is much more interesting. Glutamine is larger than leucine (double mutant) or threonine (wild-type), and is a polar residue unlike the double mutant's leucine. Figure 3-8 shows a comparison of the 41 and 124 residues of the mutants compared to WT, and it can be seen how the side chain of the glutamine is placed so as to sandwich the nicotinate ring with the flavin ring system. The NE2 nitrogen is in van der Waals contact with the nicotinate C2 (3.59 Å).

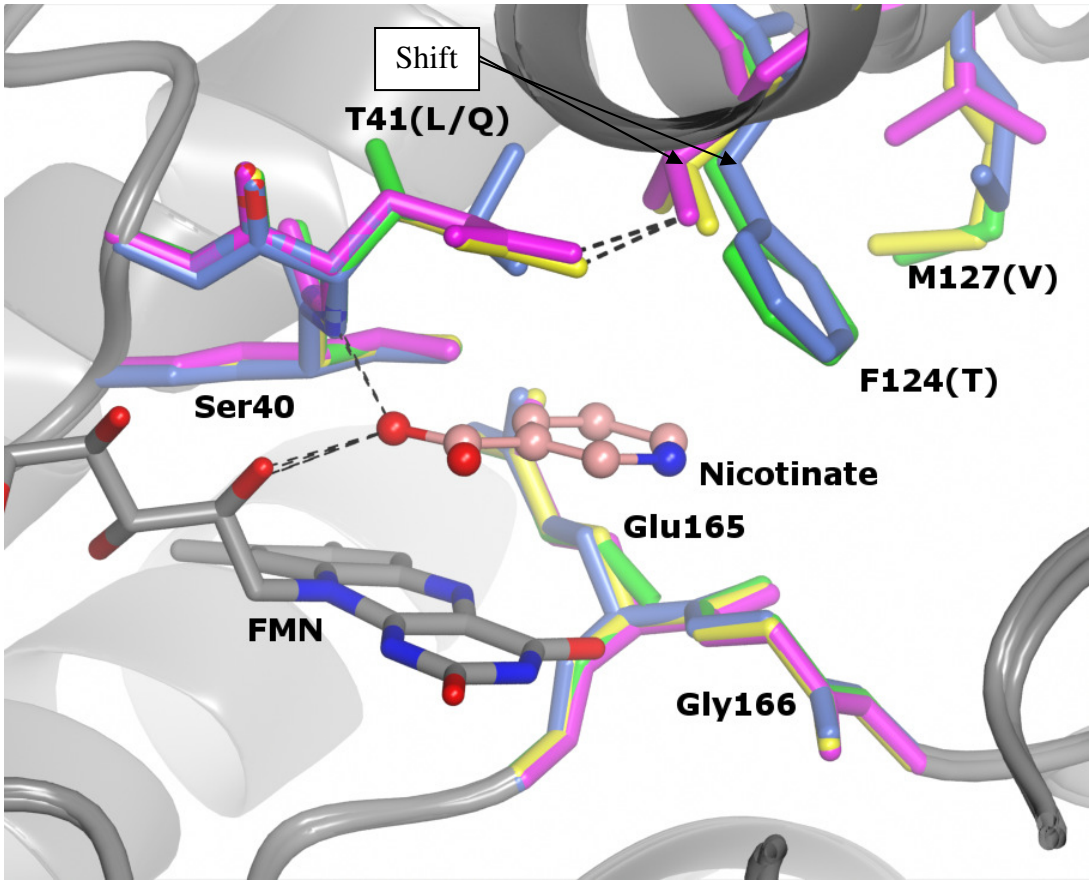


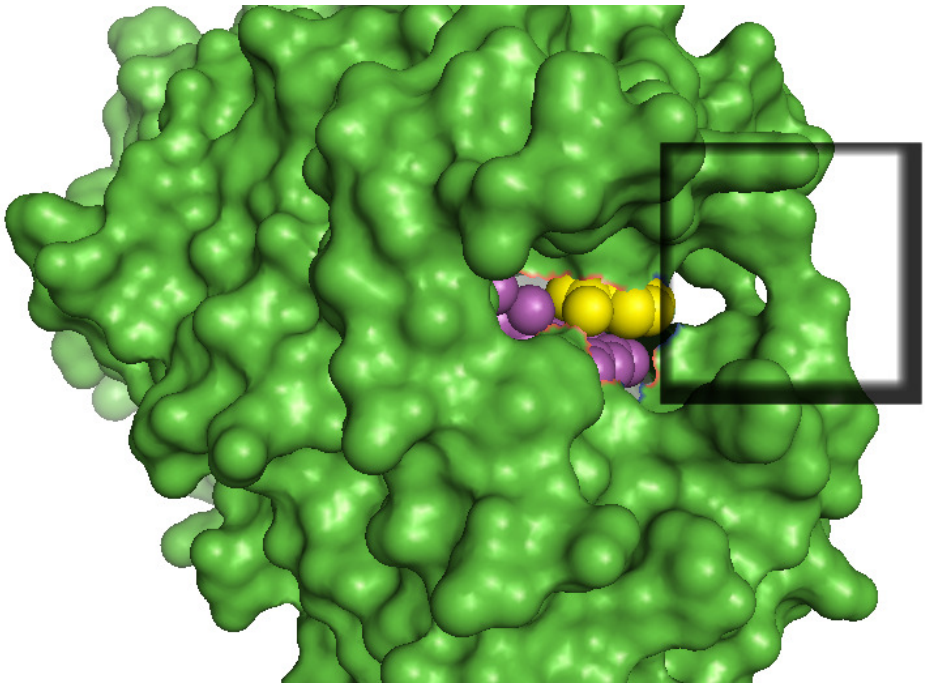
Figure 3-8: Comparison of residues 41 and 124 in T41L/N71S (blue), T41Q/N71S/F124T (yellow) and T41Q/N71S/F124T/M127V (purple) compared to wild-type NfsB (green). H-bonds are shown as black dashed lines. The shift in the position of the 124 residue in T41Q/F124T containing mutants is indicated. Figures produced using CCP4 Molecular Graphics.

Whilst the T41L mutation causes a slight shift of the Phe124 and adjacent residues away from the Leu41 (Figure 3-6), the combination of the T41Q and F124T mutations have a slight effect the other way, pulling the backbone of these residues closer to Gln41 by 0.8 Å (coordinate error - 0.22 Å) (Figure 3-8). The formation of a hydrogen bond of 3.10 Å between the amine nitrogen of Gln41 and the alcohol oxygen of Thr124 in this mutant explains this movement.

3.8.4 T41Q/N71S/F124T/M127V Quadruple Mutant Structure

The first three mutations of the quadruple mutant are the same as the triple mutant. The Ser41 residues align perfectly, and the Gln41 and Thr124 residues are also very similar. Small changes in the backbone position of the Phe124-containing loop mean that the hydrogen bond distance between Gln 41 and Thr 124 is slightly smaller in the quadruple mutant structure (2.7 Å) than the triple mutant (3.1 Å). The additional mutation in the quadruple mutant is M127V, and was poorly modelled in comparison to the others (Figure 3-5 d)). The side chain change is shown in Figure 3-7 and is too far away to have a direct effect on the active site and there are no changes seen in the other active site. The major change caused by the M127V mutation is to open up the smaller of the two channels leading to the active site. A comparison of the small channel in the triple and quadruple mutants is shown in Figure 3-9, where the first image shows the two channels, the smaller one to the right of the FMN and nicotinate and the other through which the active site is viewed. The M127V mutation causes the smaller channel to open up as shown in Figure 3-9 b). This is caused in part through replacing the methionine with a smaller valine, but this also caused a shift in the side chain position of Asn67 (shown in Figure 3-10) and results in a slight shift in the position of helix F (residues 111-130) (0.5 Å at max). There is also a shift of 1.3 Å in the position of the loop between helices F and G and one of 1.4 Å in the loop between sheet 2 and helix E. These changes have the effect of widening the channel, causing it to open up.

a)



b)

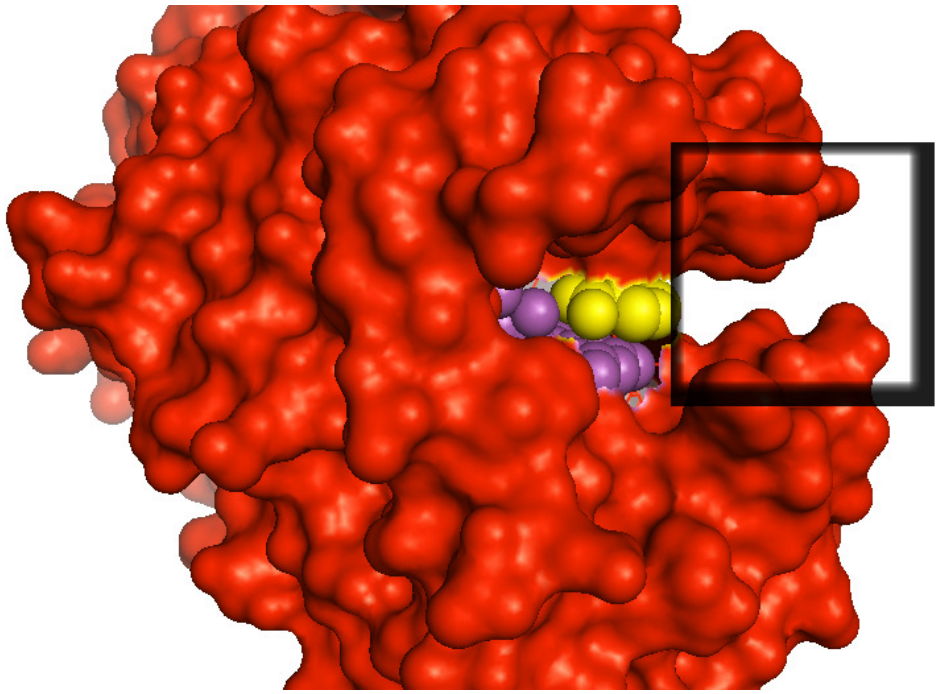


Figure 3-9: Molecular surface representations of NfsB triple mutant a) (green) and quadruple mutant b) (red). The FMN (purple) and nicotinate (yellow) molecules are represented by spheres and show the location of the active site. The two active site channels can be seen in a), the largest is the one through which one looks towards the FMN and nicotinate, the smaller is in the dark square. This smaller channel is opened up and merged with the larger in the quad mutant, b). The surface shown is the Connolly Surface (Connolly, 1983). Figures produced using PyMOL.

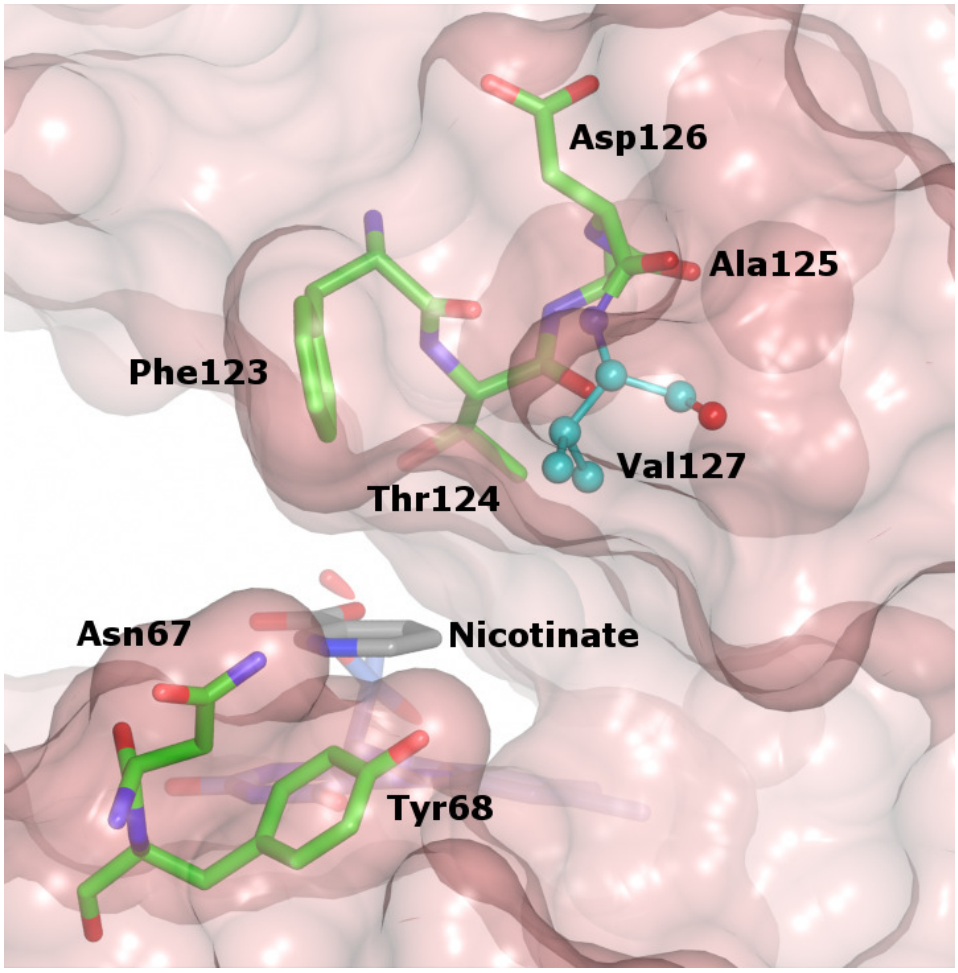


Figure 3-10: M127V mutation (light blue) in the active site of the T41Q/N71S/F124T/M127V quadruple mutant. Nearby residues are shown in green including Asn67, which is shifted to open the smaller channel to the active site (surface shown). Figure produced using CCP4 Molecular Graphics.

3.9 Ligand Soaking Attempts

Up until this point the only structures elucidated in this study contained nicotinate in the active site. In a previous study it was found that soaking a crystal in a citrate buffer first, followed by ligand solution, allowed the substrate nitrofurazone to be soaked into the active site (Race *et al.*, 2005). The same thing was attempted here. Soaks of the double and triple NfsB mutants with nitrofurazone, nitrofurantoin, menadione and metronidazole (amongst others) all proved unsuccessful. This was possibly because

the crystal soaked in the previous study to produce the nitrofurazone complex was of a monoclinic form, which did not appear in the same conditions in this study. As mentioned earlier, the only crystal form observed in the acetate buffer at pH 4.6 in this study were tetragonal plates.

In all of these soaking attempts, the active site contained some density but not enough to model a ligand confidently. Citrate has previously been used to help remove acetate from the active site of the enzyme to allow NFZ to be soaked in (Race *et al.*, 2005). Crystals of both the T41L/N71S and T41Q/N71S/F124T mutants were soaked in mother liquor but containing citrate buffer instead of acetate buffer (both 100 mM at pH 4.6) before flash freezing, to attempt to soak citrate into the active site. This gave different results for the two mutants. The citrate ligand could be modelled into the active site of the triple mutant but not of the double mutant. In the double mutant there was no longer density for the acetate directly over the flavin ring system, but rather a molecule of acetate was modelled further out of the active site, close to the side chain of Phe124. This was not in a position where it could clash with citrate if it were bound identically to the triple mutant.

3.9.1 T41Q/N71S/F124T Triple Mutant Bound to Citrate

In the citrate-bound triple mutant structure, the residues align almost perfectly with the nicotinate bound structure, with the exception of the poorly modelled and flexible Phe70. Citrate has been modelled into both active sites of the protein but in different binding modes. Both active sites are shown in Figure 3-11. The Gln41 side chain

hydrogen bonds to the alcohol group of citrate in both active sites (2.72 Å and 2.82 Å respectively). The hydrogen bond between the Gln41 and Thr124 can also be clearly seen in each active site.

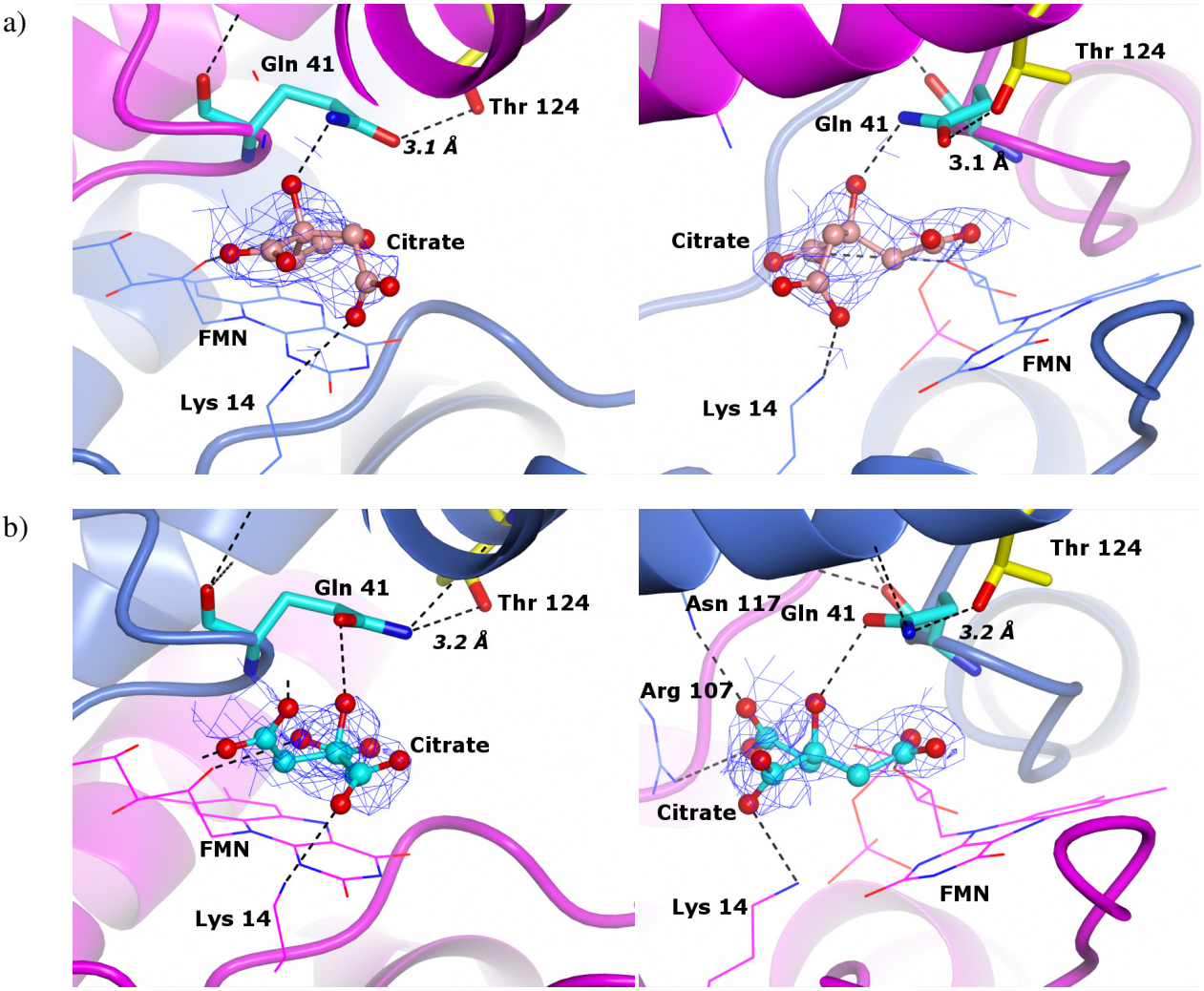


Figure 3-11: Electron density of citrate bound in the active site of NfsB T41Q/N71S/F124T mutant. The two active sites are shown in a) (pink citrate) and b) (light blue citrate); each with two images 90 ° apart. The two monomers are coloured purple and blue (ribbons and carbon atoms). Ball and stick diagrams have been used for the citrate. The Gln41 (light blue) and Thr124 (yellow) residues are shown as cylinders and the H-bond (dashed line) between is annotated. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.0 σ) has been clipped to within 2.5 Å of the citrate. Figure produced using CCP4 Molecular Graphics.

The citrate is primarily bound to the active site through hydrogen bonding to protein and FMN. The pattern of hydrogen bonding for each active site is shown in Figure 3-12. The citrates in the two different sites effectively have the central carbon atom and alcohol group in similar positions, and differ in terms of which of three sites (labelled 'A-C' in Figure 3-12) contains the shorter chain carboxylate. Site A is close to the amine or Lys14, site B is over the flavin ring system, and site C is more towards the phosphate binding pocket. In the first active site (Figure 3-12 a)) the short carboxylate chain is in position C, while in the second (Figure 3-12 b)) it is in position A. Due to the different lengths of the arms of the citrate (two contain methylene groups), the change in conformation impacts on the overall position of the citrate. At pH 4.6, two of the carboxylic acid groups of citrate should be protonated and one should be negatively charged. For the first active site it is likely that the carboxylate in the 'A' position in Figure 3-12 a) will form a salt bridge with the positively charged Lys14 amine. For the second active site the situation is less clear as the distance between the Lys14 amine and the carboxylate is greater (3.1 Å vs. 2.6 Å), though since it is the only residue binding the citrate that is positively charged it is still likely that that the carboxylate in position 'A' would be negatively charged although the pK_a may be rather different in the protein.

a)

b)

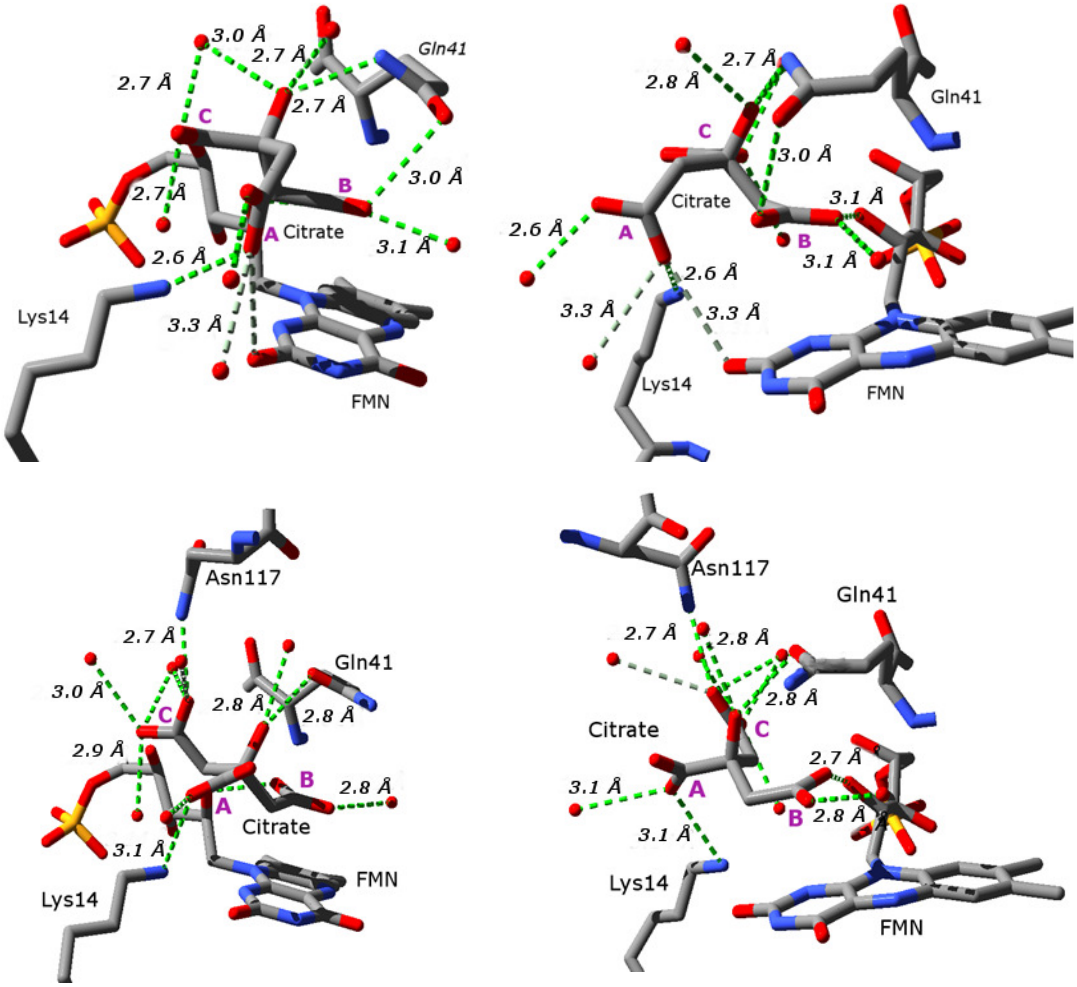


Figure 3-12: Hydrogen bonding pattern of citrate in NfsB T41Q/N71S/F124T. Each active site is shown, a) and b), with images separated by 90°. Residues and water molecules are shown that hydrogen bond to the citrate ligand. The three possible positions of the carboxylic acid groups are labelled A-C. Figure produced using Swiss-PdbViewer and POV-Ray.

The binding pattern of citrate would not be possible without the T41Q mutation which might explain why it was not found in the double mutant. Citrate was tested as an inhibitor at neutral pH, but showed no significant inhibition at 25 mM. Above this concentration the effect of raising the ionic strength begins to dominate the interactions, making testing inhibition difficult. However the crystals were grown at pH 4.6 (measured with pH paper), where only one of the carboxylic acid groups would be expected to be deprotonated in solution (although pK_a's can be much different in proteins). Since the binding shown is primarily through hydrogen bonding to the protonated carboxyl groups, this would not be possible at pH 7 and so it is plausible that this would explain why no inhibition was seen.

3.9.2 Double Mutant Empty Form and Triple Mutant Acetate Bound

Attempts to soak citrate into the double mutant produced an empty active site (Figure 3-13 – green). There was a molecule of acetate bound but it was not in the same position as any previous acetate bound structures (Race *et al.*, 2005). This acetate was not bound above the aromatic ring system but was found closer to the Phe124 residue. The density above the flavin ring system was less clear, a single water molecule was modelled in but it is probable that the citrate had caused the acetate to be removed from the active site but did not bind strongly enough to be seen in the electron density. So the electron density in the active site was caused by a mixture of citrate, acetate and solvent. The fact that citrate was seen with the triple mutant and not the double mutant may be explained by the hydrogen bond from the triple mutant Gln 41 residue to the citrate (Figure 3-12). Citrate may still bind the double mutant as acetate is seen in the

old position without changing the buffer to citrate (data not shown), but would lack this interaction as there is nothing in place of the amine atoms from Gln 41.

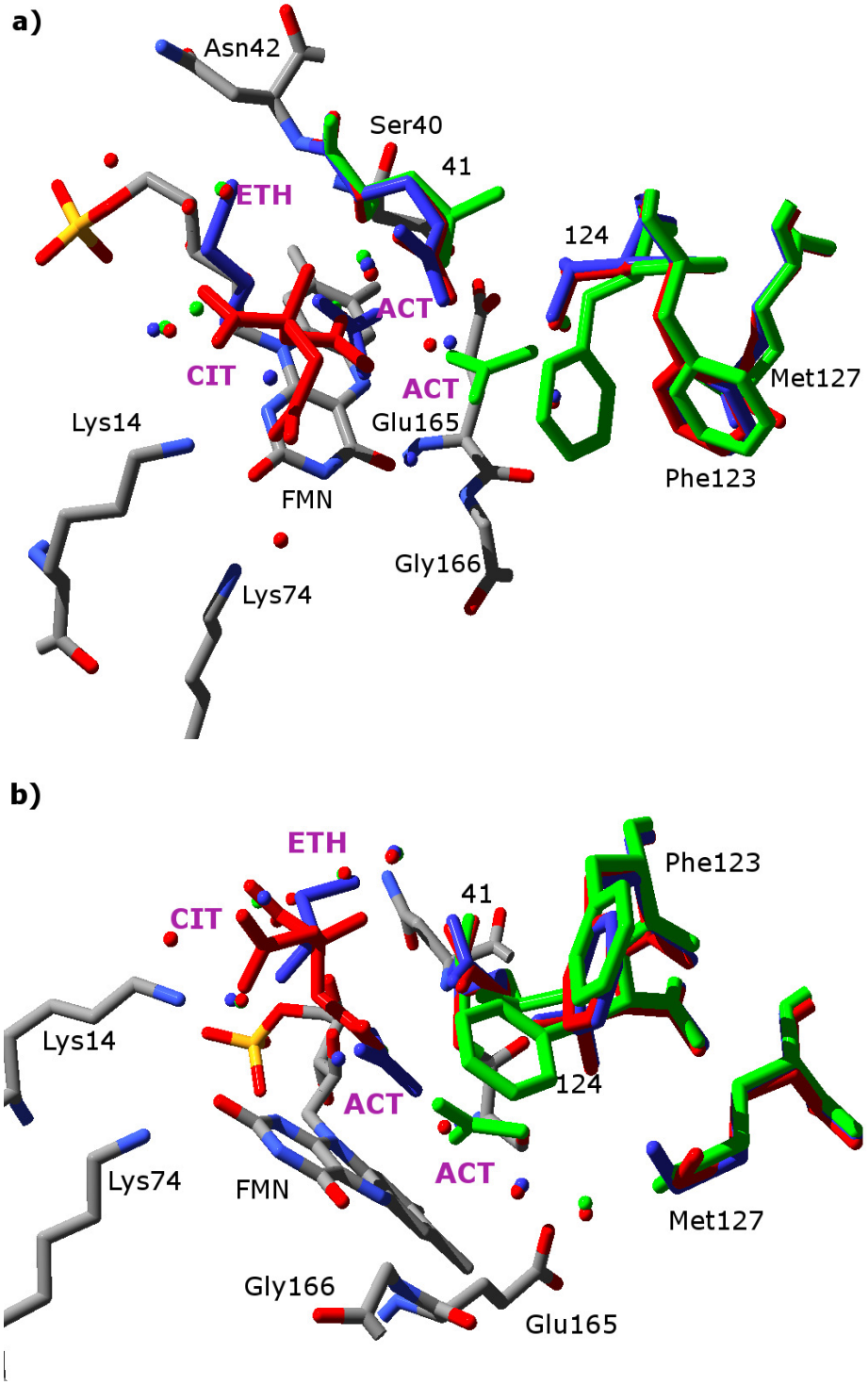


Figure 3-13: Comparison of the active site of the NfsB T41L/N71S double mutant empty active site (green), and the T41Q/N71S/F124T triple mutant bound to acetate (blue) and citrate (red). Only residues which vary are shown for all the mutants. Ligands are labelled in purple, citrate (CIT), acetate (ACT) and ethylene glycol (ETH). Water molecules are shown as spheres. a) and b) are shown 90° apart. Figure produced using Swiss-PdbViewer and POV-Ray.

For the triple mutant, a complex with acetate was determined from a crystal in acetate buffer, in the absence of nicotinate. Acetate is bound in the triple mutant structure in a different conformation than had been seen with previous structures. It is rotated around 120° and is further away from the flavin ring system. There is a hydrogen-bond from the acetate to the side chain of Gln41.

3.10 Attempts to Co-Crystallize with Prodrugs

With crystals obtained at neutral pH, the decision was made to attempt co-crystallization with prodrugs. Soaks were not possible due to the fragile nature of the crystals. Co-crystallization with the dinitrobenzamide prodrugs (the 'SN-' prodrugs in Figure 2-3) led to the formation of new crystal forms, some larger and better diffracting than the previous ones (Section 3.4). One previously-unseen crystal form led to a structures at higher resolution than has previously been reported at 1.08 Å (the previous highest being 1.7 Å (Lovering *et al.*, 2001)). With all the dinitrobenzamide prodrugs (including CB1954) a limiting concentration of prodrug was seen, above which no crystals formed. The datasets presented here are with the highest concentrations of prodrug that still allowed crystals to form.

With the crystals that had been co-crystallized with CB1954 and each of the other dinitrobenzamide prodrugs detailed in Figure 2-3, the active site contained definite density above the flavin ring system. This density was nearly identical for all the prodrugs, but was not seen with lower concentrations of the prodrugs or in their absence. It matched a 6-membered ring with a single substituent (Figure 3-14). All the

prodrugs should have four ring substituents, but even at low contour levels of 0.4 σ there are no peaks in the electron density map allowing the substituents to be located. Since for CB1954 all substituents are of a similar size, the possibility was considered that the molecule was in multiple orientations and angles, however there should always be a substituent opposite the one that is seen. Also the same density was seen with the 'SN' prodrugs, implying that it would be unlikely to be due to a contaminant in the prodrug solution. Since the solvents were different for CB1954, which was in a 2:7 solution of NMP:PEG 300, and the 'SN' prodrugs, which were in acetonitrille, it is unlikely that a contaminant in the solvent could have caused the density. The ring and substituent observed in the structures are identical in shape to nicotinate, and are found in a nearly identical position to nicotinate in the structure. However, these crystals were never in the presence of nicotinate. It is possible that the prodrugs have been degraded in the solution or X-ray beam, causing all but one of the substituents to react and be removed.

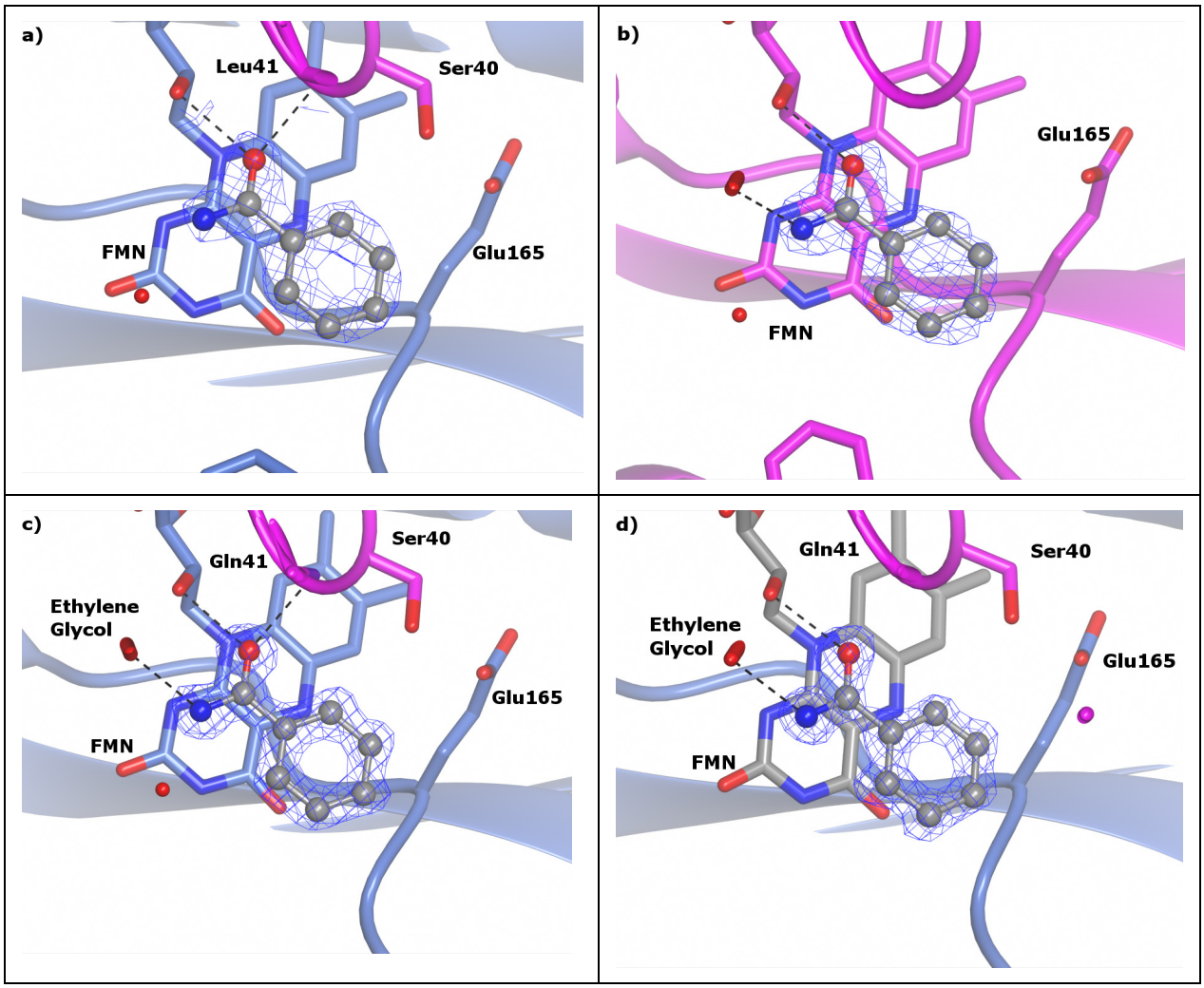


Figure 3-14: Examples of the unbiased electron density in the active sites of NfsB mutant crystals co-crystallized with prodrugs. Figure a) Double CB1954, b) Double SN27683, c) Triple SN27683 and d) Triple SN27686. The model of a benzene ring and single substituent (amide) is shown as grey ball and sticks, modelled into the density. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

3.11 Analysis of New Crystal Forms

Four different crystal forms of NfsB have previously been observed (Lovering *et al.*, 2001, Race *et al.*, 2005). In this study are presented crystal structures from two additional, previously unseen crystal forms. These are all summarised in Table 3-5. The two new forms are both monoclinic, containing two copies of the NfsB monomer

in the ASU. Whether it was the presence of the prodrug or the acetonitrile (used as a solvent) that prompted the growth of these new crystal forms is unclear. However the new forms did not appear when acetonitrile was used on its own. The Monoclinic 3 form has a similar solvent content to the tetragonal (both ~ 45-46 %), but the Monoclinic 2 form is more tightly packed having a solvent content of around 41 %.

Identifier	Space Group	a (Å)	b (Å)	c (Å)	β (°)	Copies in ASU	Example
Tetragonal	P 4_1 2_1 2	55	55	260	-	2	pdb I.D. 1ICR (Lovering *et al.*, 2001)
Orthorhombic 1	P 2_1 2_1 2_1	55	120	145	-	4	pdb I.D. 1ICV (Lovering *et al.*, 2001)
Orthorhombic 2	P 2_1 2_1 2	155	45	55	-	2	pdb I.D. 1YLU (Race *et al.*, 2005)
Monoclinic 1	P 1 2_1 1	70	55	115	105	4	pdb I.D. 1ICU (Lovering *et al.*, 2001)
New forms:							
Monoclinic 2	P 1 2_1 1	55	60	75	100	2	Triple SN27683
Monoclinic 3	C 1 2 1	115	45	95	115	2	Double SN27683

Table 3-5: Summary of the crystal forms seen with NfsB and natural amino acid mutants. For previously seen forms examples are given from the literature, for the new forms examples are given from this study. Average cell dimensions are given to the nearest 5 Å as there was up to 5 Å variation in cell dimensions in crystals of the same form.

3.12 Stopped-Flow Kinetics

To analyse the rate of reduction and K_d of NADH independently of the second substrate, stopped flow kinetic experiments were performed. The nature of the bi-bi mechanism of catalysis by the nitroreductase enzymes is such that it is formed of two half-reactions (labelled a) and b) on Figure 1-17). The steady-state kinetics presented in this study monitor this reaction overall, but not the steps in isolation. The individual steps can be studied by employing stopped-flow kinetic measurements. Stopped-flow kinetic data for the reductive half reaction, step b) in Figure 1-17, of WT-NfsB, F124N, T41L/N71S and T41Q/N71S/F124T with CB1954 and nitrofurazone has been published (Jarrom *et al.*, 2009). The oxidative half-reaction, the reduction of NADH, step a) in Figure 1-17, as catalysed by WT-NfsB, T41L/N71S and T41Q/N71S/F124T, is described here.

The oxidative half reaction was followed by monitoring the change in absorbance of the tightly bound FMN cofactor at 454 nm. Data sets were collected over a period of 0.05 s, and data before 1 ms was not including in fitting. The averaged traces for the NfsB T41L/N71S double mutant are shown in Figure 3-15. The exponential decay gets steeper at higher concentrations of NADH. An oscillation was observed, at approximately 50 Hz, which was consistent across all traces, implying interference, possibly electrical. However all the traces fit well to Equation 2-13, with high t values (all > 7 and detailed in Table 3-6) and P values of <0.0001.

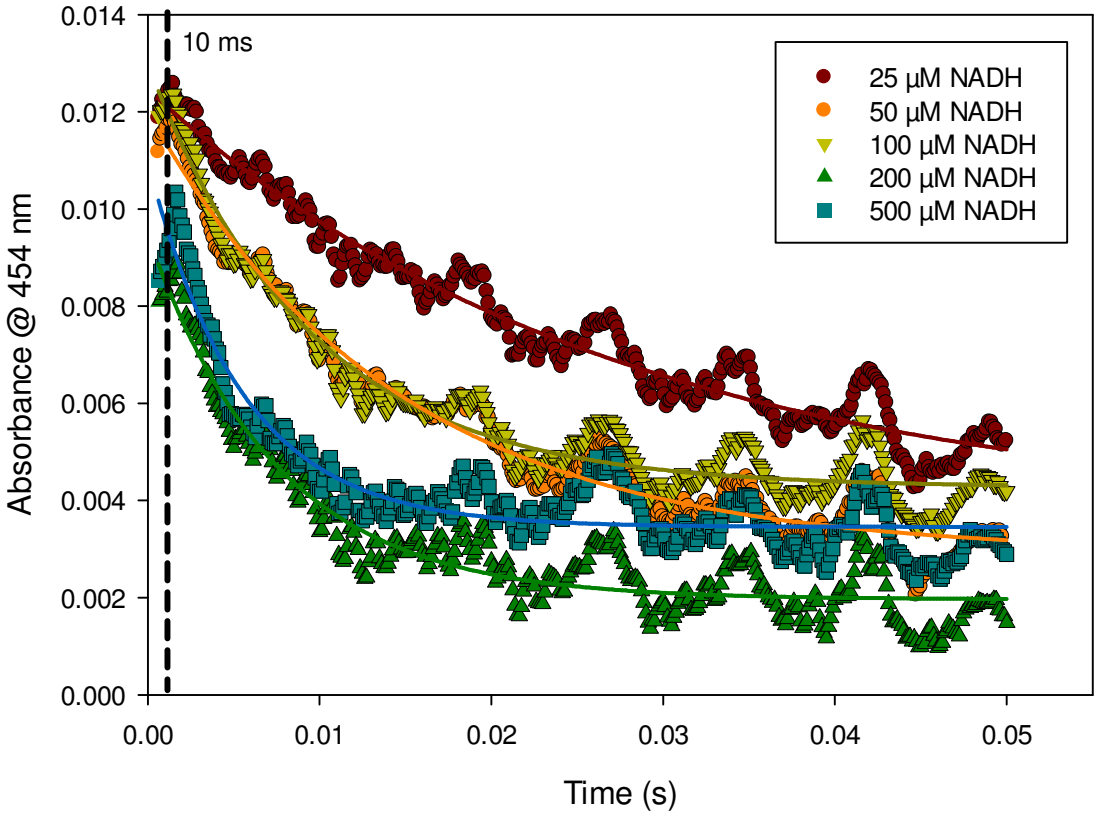


Figure 3-15: Stopped-flow kinetics traces for the reduction of FMN bound to the NfsB T41L/N71S double mutant by varying final concentrations of NADH. Data post 10 ms, indicated by the dashed line, was fitted to Equation 2-10 giving the lines shown.

The rate constants for each run were plotted against the concentration of NADH, and this was done for NfsB WT, T41L/N71S and T41QN71SF124T (Figure 3-16). The data was fitted to a hyperbola, either to Equation 2-15, to determine k_2 (the rate of oxidation of NADH by the NfsB bound FMN, see Figure 2-8) and K_d, or to Equation 2-16 to determine k_2/K_d, the apparent second order rate constant.

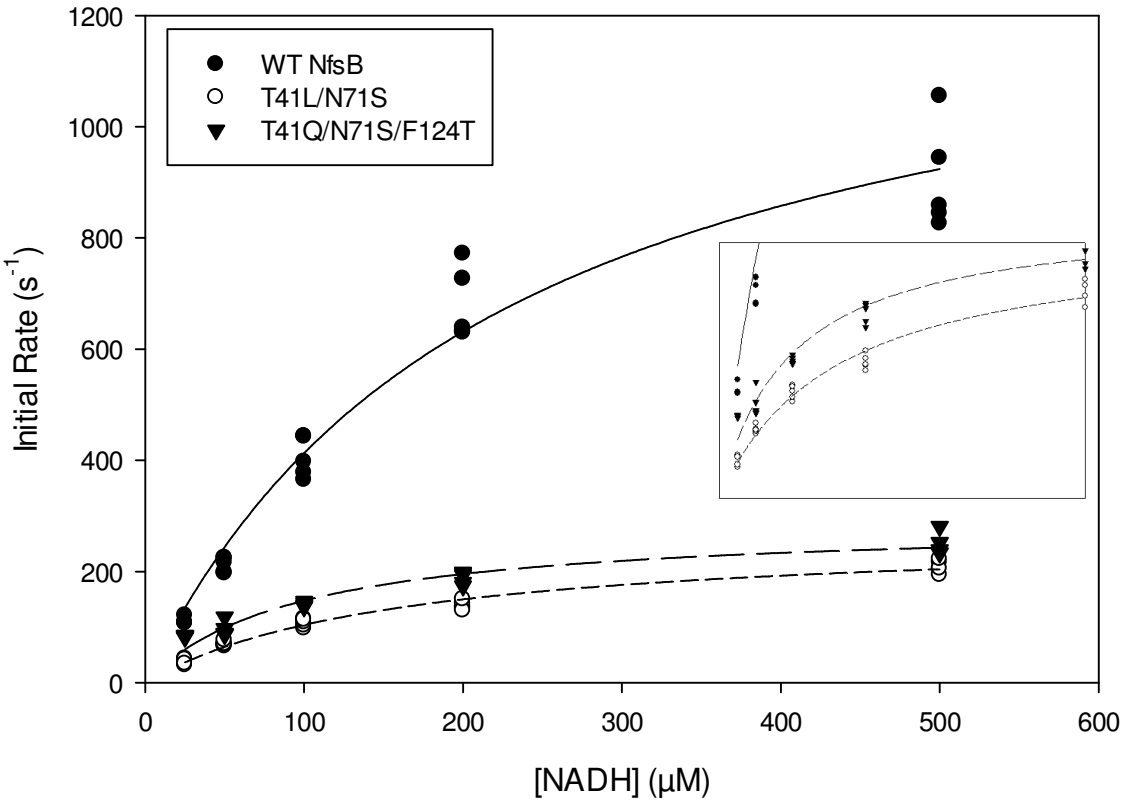


Figure 3-16: NADH oxidation rates of WT, double and triple NfsB mutants as measured by stopped-flow kinetics. The insert shows a close up of the double and triple mutant data, showing that the hyperbolas for each do indeed level off to a maximum.

The kinetic parameters for the oxidation of NADH by NfsB WT, T41L/N71S and T41Q/N71S/F124T are shown in Table 3-6. The apparent second order rate constants (k_2/K_d) should be comparable with the specificity constants (k_{cat}/K_m) for NADH, previously determined through steady-state kinetics (Jarrom *et al.*, 2009), although they all appear to be roughly twice as large. The values of both k_2/K_d and k_{cat}/K_{mNADH} are approximately 2-fold larger for Wild-type than both the double and triple mutants. This could be due to accuracies in assessing protein concentrations for the steady-state, or due to the extinction coefficient being incorrect. Thus when comparing both

mutants to WT it becomes apparent, that as mutants have been selected for increased activity with CB1954, they have lowered the specificity constant for NADH. Comparing the steady-state rate constants (k_{cat}) with the k_2 rate constants gives a clue as to the rate-determining step. For WT-NfsB, k_{cat} is well below k_2 implying that the oxidation of NADH is not the rate determining step. The maximum rates seen for the reduction of CB1954 in WT, double and triple mutants are also higher than the steady-state k_{cat} values (Jarrom *et al.*, 2009). This implies that the rate limiting step is the release of products. However for the double and triple mutants we can now see that the difference between k_{cat} and k_2 is much less, and if this trend were to continue with mutants developed in the future, the rate limiting step may change.

	k_2 (s^{-1})	t	P	K_d (μM)	t	P	k_2/K_d (s^{-1} μM^{-1})	t	P	k_{cat}/K_{mNADH} (s^{-1} μM^{-1}) (Jarrom *et al.*, 2009)	k_{cat} (s^{-1})	K_{mNADH} (μM)
WT-NfsB	1340 ± 84	16	<0.0001	224 ± 30	7.4	<0.0001	5.98 ± 0.47	13	<0.0001	3.46 ± 0.6	140 ± 32	40 ± 12
T41L/N71S	269 ± 11	25	<0.0001	158 ± 15	11	<0.0001	1.70 ± 0.098	17	<0.0001	0.62 ± 0.08	153 ± 8	253 ± 41
T41Q/N71S/F124T	290 ± 13	22	<0.0001	96.2 ± 12	8.1	<0.0001	3.01 ± 0.26	12	<0.0001	1.33 ± 0.077	181 ± 7	136 ± 12

Table 3-6: Kinetic Parameters for the oxidation of NADH by NfsB WT, T41L/N71S and T41Q/N71S/F124T as determined by stopped-flow kinetics. The specificity constants (k_{cat}/K_m) are taken from kinetic parameters for CB1954 reduction (Jarrom *et al.*, 2009).

3.13 Summary

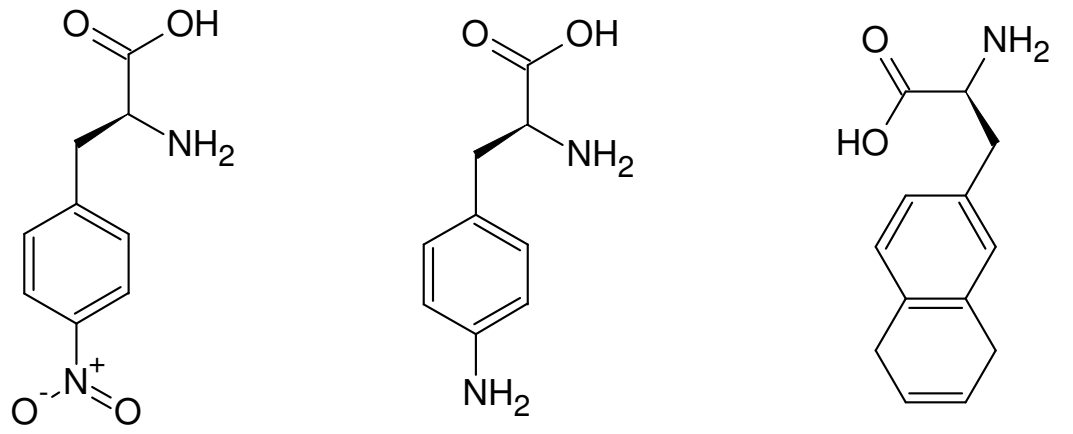
The crystal structures of the double, triple and quadruple mutants of NfsB that show the greatest improvement of WT for the reduction of CB1954 have been determined. Each was determined bound to the NAD(P)H analogue, nicotinate, and the differences in the side chain positions give insights into the improvements seen in kinetics. Attempts to solve prodrug bound structures of each mutant were unsuccessful, although the density present in these experiments suggested that the drug was being degraded. Stopped-flow kinetic studies have also been used to analyse the oxidation of NADH by WT, double and triple mutant NfsB, indicating that as the mutants have been selected for their ability to reduce CB1954, they have become less active for NADH.

The improvements in k_{cat}/K_m in the mutants over wild-type could be caused by two factors. These improvements are either caused by altering the properties of the FMN, such as in the case of the N71S mutation, which has been shown to alter the redox potential of the FMN (unpublished work by Dr Ross Anderson, University of Bristol), or by changing the ligand binding site. The other mutations would appear to primarily affect the ligand binding site, they improving the affinity of the enzymes for CB1954.

4. NfsB Unnatural Amino Acid Mutants

4.1 Incorporation of Unnatural Amino Acids

The activity of NfsB with CB1954 has been shown to be improved upon incorporation of several unnatural amino acids residues in place of the Phe124 residue (Jackson *et al.*, 2006). The most active of these unnatural amino acid mutants were *para*-nitrophenylalanine (pNF), *para*-aminophenylalanine (pAF), and naphthylalanine (Nap) (Figure 4-1). Dr Ryan Mehl (Oregon State University, USA) provided the plasmids encoding these mutants, for the purpose of crystallographic studies. Different plasmids and purification methods were used from those described previously for the NfsB mutants containing natural amino acids, and the purifications were performed exploiting an N-terminal His6-tag (described in Section 1.6.2.3.4).



p-Nitrophenylalanine *p*-Aminophenylalanine Naphthylalanine

Figure 4-1: Structures of the unnatural amino acids used in this study. Figure produced using ChemSketch (ACD/labs).

4.2 Growth and Induction of Unnatural Amino Acid Mutants

Small scale (50 ml) cultures of auto-induction media (see Section 2.3.2) with WT-NfsB and the three unnatural amino acid mutants were made to test growth and induction. Cell growth was monitored by following turbidity at 600 nm. Cell growth peaked at around 22 hours post-inoculation. Growth of *E. coli* expressing WT NfsB peaked at between 8-9 OD units, whilst cell growth for *E. coli* in the presence of unnatural amino acids peaked at around 5-6 OD units. NfsB expression was monitored by SDS-PAGE (with samples all having the same OD, as described in Section 2.3.2) and peaked by 22 hours in all cases (Figure 4-2). The WT (not shown) and pNF showed the highest levels of NfsB expression, while the Nap mutant was very poorly expressed.

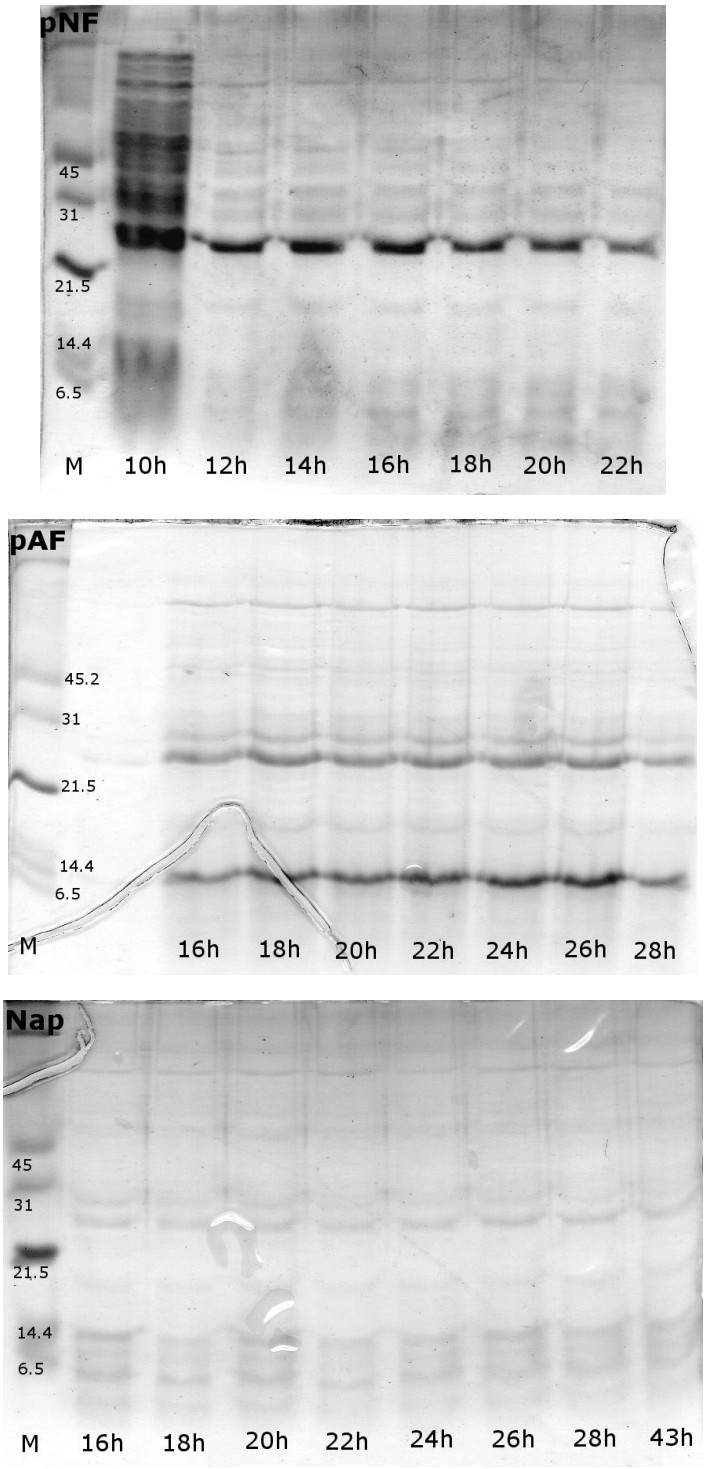


Figure 4-2: SDS-PAGE analysis of protein expression. Images are labelled by unnatural amino acid at position 124 in NfsB. Time of sample is indicated for each lane in hours.

4.3 Purification of His-tagged NfsB

Unlike previous NfsB and mutant proteins purified by this group (including those detailed in Chapter 3), the unnatural amino acid mutants were purified by means of an N-terminal His-tag. They were purified first with a Nickel-NTA column, eluting with increasing imidazole (Figure 4-3), and then, since further purification was required, the protein was purified on a Phenyl-Sepharose column as for the natural NfsB mutants (Figure 4-4). Figure 4-5 shows SDS-PAGE analysis of the final protein fractions used in this chapter. A poor yield of the Nap mutant (see expression experiment in Figure 4-2) meant that there was insufficient protein for a single column. This did not purify the protein sufficiently to allow quantitive analysis. Repeated growth experiments (varying conditions including concentration of components and addition of glucose) showed poor expressions of the Nap mutant and so no subsequent experiments were performed.

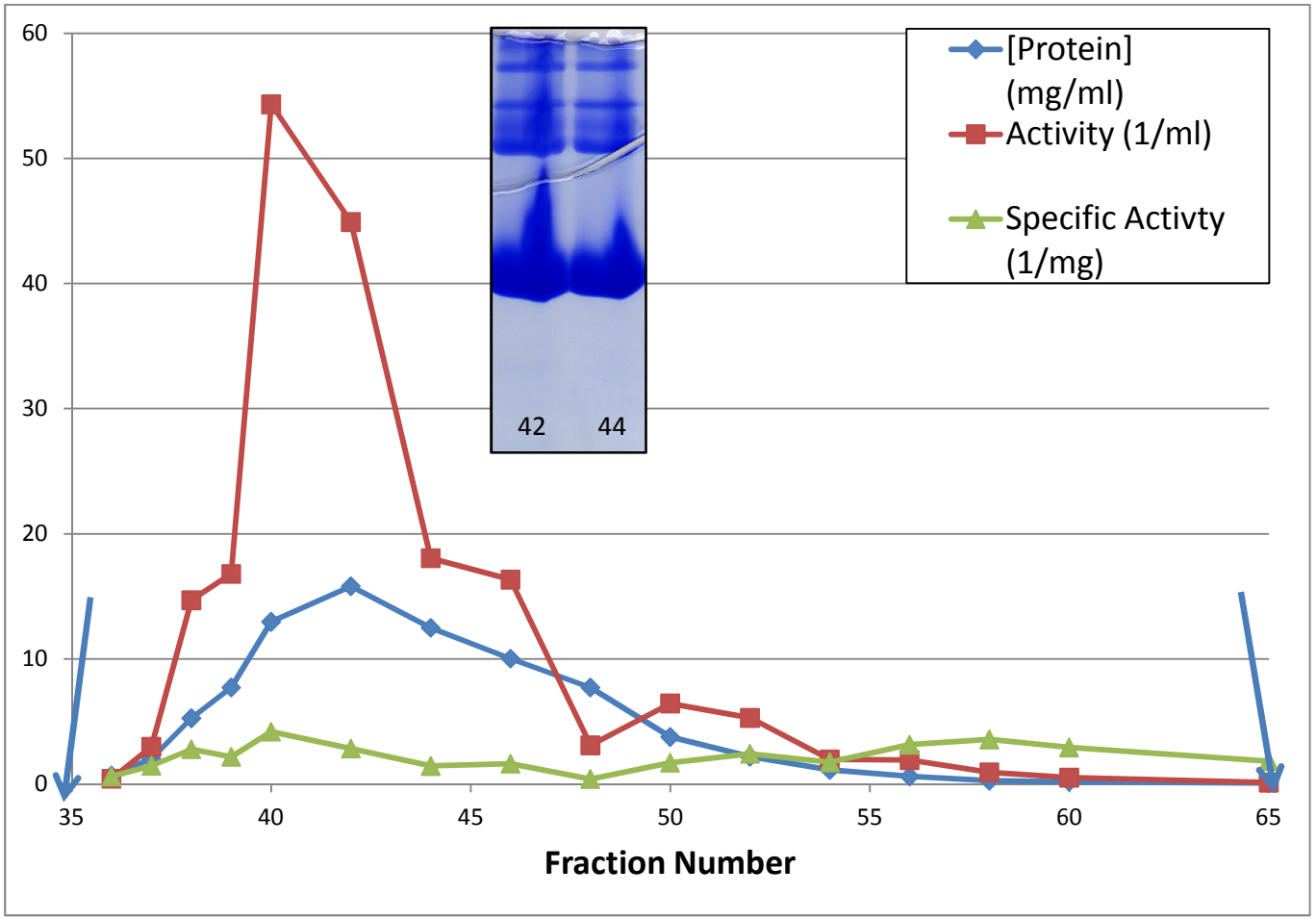


Figure 4-3: Analysis of WT-NfsB protein fractions eluted off Nickel column (see section 2.3.2, insert - SDS-PAGE). Fractions size was 2 ml and column size was 10 ml. The blue arrows show the start and finish points of the imidazole gradient run from 20 mM to 250 mM Imidazole.

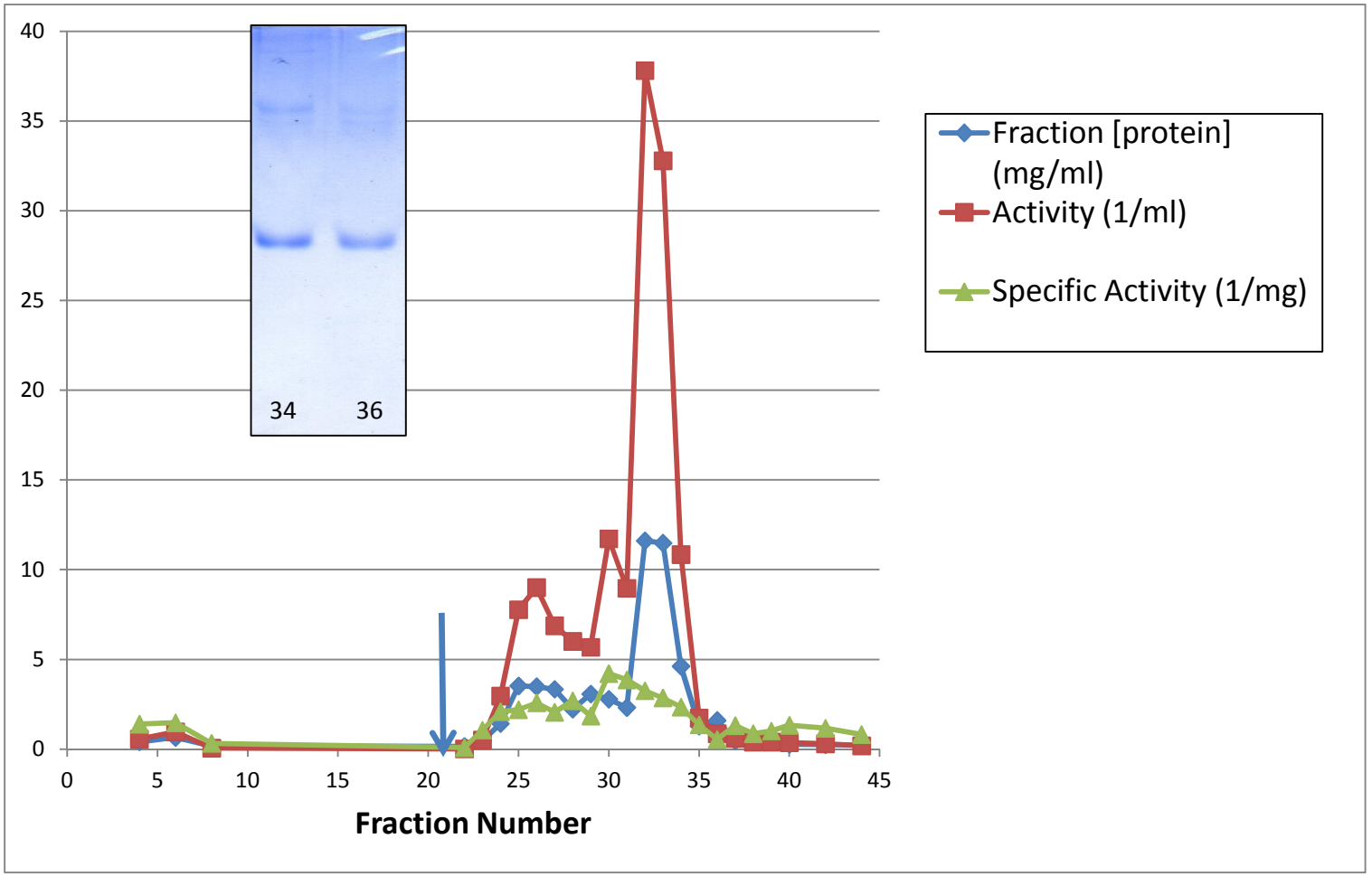


Figure 4-4: Analysis of WT-NfsB protein fractions eluted from Phenyl-Sepharose column (insert - SDS-PAGE). Fraction size was 2ml and column size was 10 ml. The blue arrow shows where the elution buffer was applied to the column.

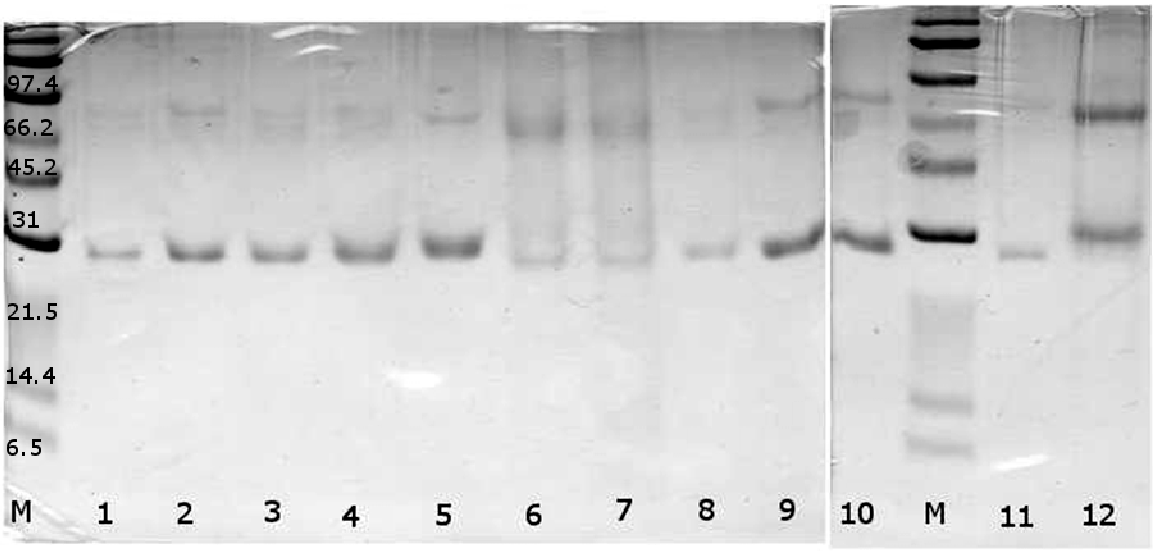
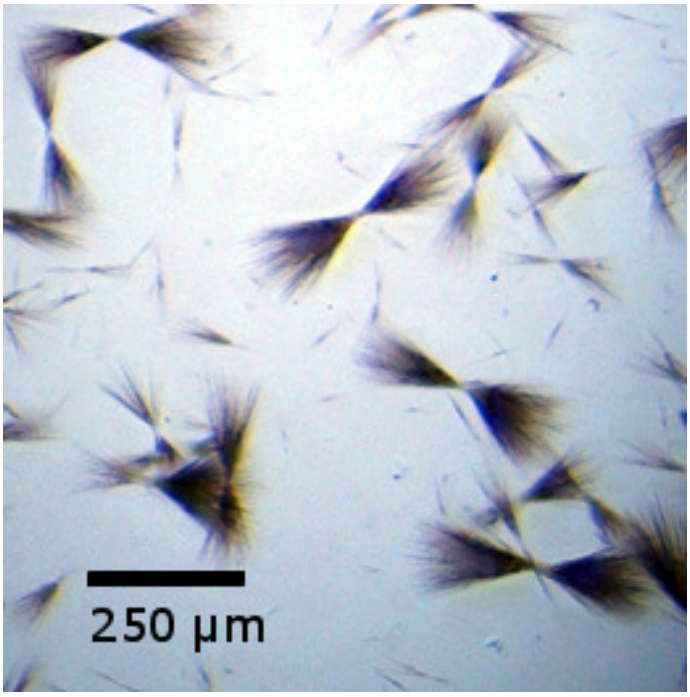


Figure 4-5: SDS-PAGE of final pooled protein fractions purified with His-tag. M – Marker, lanes 1-5 + 12 - WT-NfsB, 6-7 - Nap mutant, 8 – pAF mutant, and 9-11 – pNF mutant. Each separate lane is of a separate pool.

4.4 Initial crystal growth

The first attempts to crystallize His-tagged NfsB were based on the same conditions as those previously used for natural mutants and WT protein (Section 3.3) (Parkinson *et al.*, 2000, Lovering *et al.*, 2001). In both of these sets of conditions only tiny needle-shaped crystals growing in clusters were found (Figure 4-6). These crystals were not suitable for X-ray diffraction studies, and since this form was the only one seen (in conditions that usually gave more suitable crystal forms) the decision was made to cleave the His-tag and then attempt to crystallize again.

a)



b)

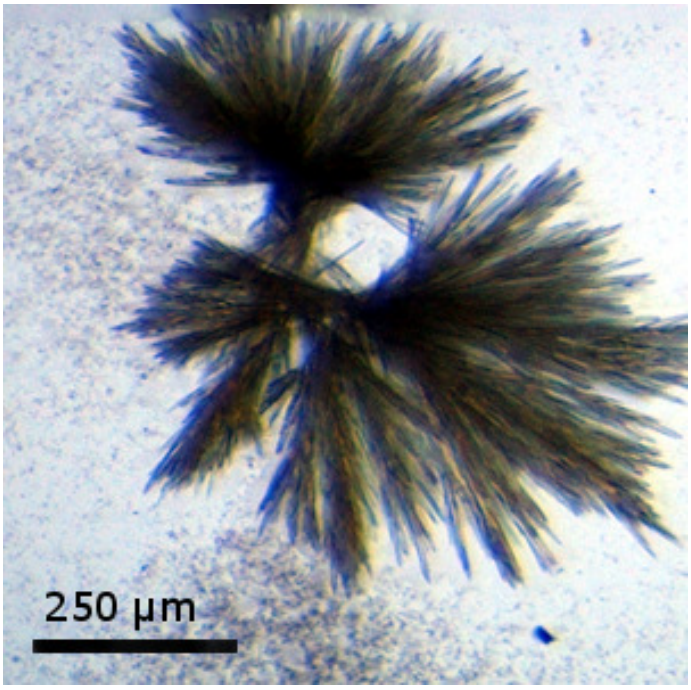


Figure 4-6: Initial crystal growth of His-tagged NfsB. a) Whole well of crystal tray showing multiple groups of thin needles, b) a close up of a single cluster.

4.5 Cleavage of His-tag

The N-terminal 6 His-tag was cleaved using a recombinant enterokinase (rEK). A test cleavage was performed with 0.02 – 0.1 U rEK in 50 µl and 0.2 µg/ml NfsB over 48 hours. The cleavage tests are shown in Figure 4-7. Once ideal conditions were identified (detailed in Section 2.3.3) further tests were performed, including for the pNF and pAF mutants at the proportions identified as best. Beyond 48 hours secondary digestion was seen in some tests (low molecular weight bands on the gel, but at 48 hours most of the His-tag cleavage had been performed. So for large scale cleavage the reaction was halted at 48 hours with 10 µg/ml PMSF (and 1 % iso-propanol as a solvent).

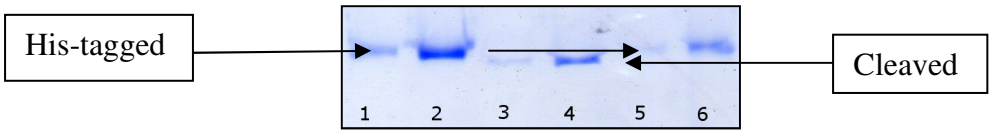


Figure 4-7: SDS-PAGE showing the digestion of the His-tag on 0.2 µg/ml WT-NfsB by rEK after 48 hours. Lanes 1-2 show His-tagged protein, 3-4 NfsB plus 0.1 U rEK, 5-6 NfsB plus 0.02 U rEK (all in total 50 µl solutions). Lanes 1, 3 and 5 were diluted 1/10 in 10 mM Tris pH 7.0. The bands in lanes 3 and 4 are lower showing that in these conditions the His-tag has been successfully cleaved.

4.6 Further crystal trials

Crystals were grown as described in Section 2.5.1.2, and appeared within 3-5 days and reached full size within a week. Thus these crystals grew more slowly than the natural amino acid mutants. Crystals of the pAF mutant grew in the presence of 15 mM nicotinic acid, whilst crystals of the pNF mutant did not. Several different crystal forms were seen, the first (Figure 4-8 a)) was a long flat needle form, mostly

multinucleated. These crystals only diffracted to around 8.5 Å, and showed high levels of mosaicity, which made indexing of diffraction patterns impossible. Another crystal form found was a variation on a monoclinic form previously observed for NfsB natural mutants (Lovering *et al.*, 2001). The individual crystals (Figure 4-8 d)), were mostly found in groups shaped somewhat like a flower bud where the individual crystals were like petals (Figure 4-8 b) and c)). In order to mount these crystals the clusters were broken up and individual crystal plates were selected; this was difficult as the crystals were very fragile. It was this crystal form that gave rise to the two structures detailed in this chapter; however the two crystals (of the different mutants in the different conditions) were in different space groups. There were other crystal forms not shown, that were mostly multinucleated and diffracted poorly.

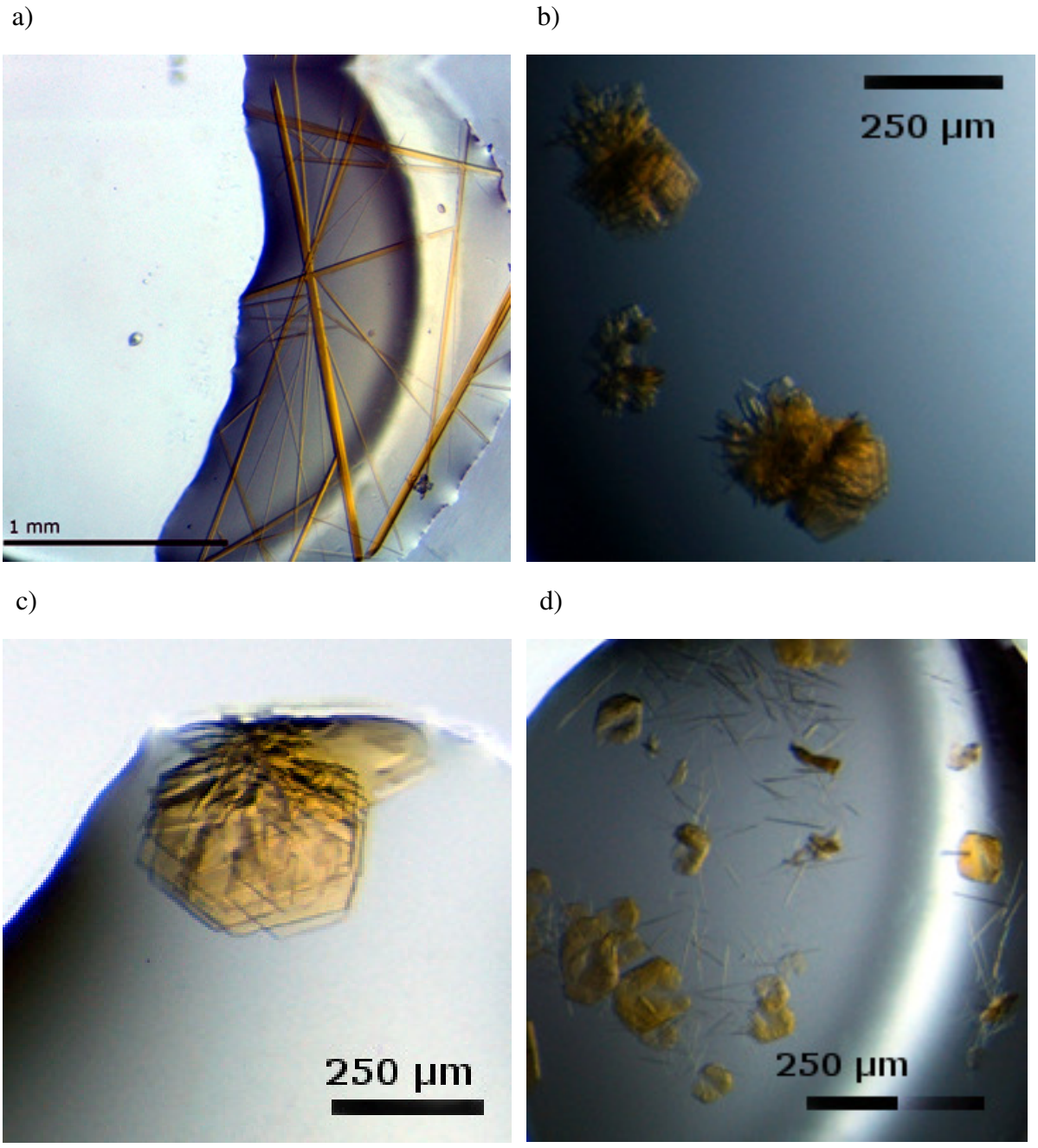


Figure 4-8: Crystals of His-tagged cleaved NfsB. a) pNF mutant crystals, large needles, b) pNF mutant crystals, monoclinic clusters, c) pAF mutant crystals, monoclinic clusters and d) pAF mutant crystals, monoclinic and needles.

4.7 Summary of Datasets

Two datasets are presented here, one of the pNF mutant with bound acetate, and another with the pAF mutant with bound nicotinate. The collection statistics are shown in Table 4-1, processing statistics in Table 4-3 and refinement statistics in Table 4-4. The pAF nicotinate structure is in a monoclinic crystal form (pdb I.D. 1ICU) observed previously (Lovering *et al.*, 2001), however the pNF acetate structure is in a crystal form never before seen with NfsB. Both diffracted to a similar resolution (~2.7 Å), which is poorer than seen with all previously observed crystal forms but enough to see any major active site differences. Unlike all the natural amino acid mutant structures presented in Chapter 3, these forms do not contain two copies of the monomer in the ASU. The pNF-Acetate structure contained 8 copies of the monomer, whilst the pAF-Nicotinate contained 4 (as did previous structures in this space group). The five amino acid remains of the cleaved His-tag (see Table 2-1 for details) were not seen in either structure.

	Crystallization			Soak
Dataset Identifier	*Precipitant – [PEG 4,000]*	*Additive*	*Protein Conc. (mg/ml)*	*Cryo. Prot.*
pNF Acetate	12%	-	11	15% Ethylene Glycol
pAF Nicotinate	13%	15 mM Nicotinic Acid	14	15% Ethylene Glycol

Table 4-1: Crystallization conditions for the crystals giving rise to the NfsB unnatural amino acid mutant crystal structures detailed in this chapter. All crystals were grown via a hanging drop method with 100 mM Sodium Acetate pH 4.6

	General					Unit Cell				
Dataset Identifier	*Active Site Contents*	*Collected at*	*Beamline*	*Wavelength (Å)*	*Degrees of Data (°)*	*Space Group*	*a (Å)*	*b (Å)*	*c (Å)*	*β (°)*
pNF Acetate	Acetate	Diamond	I04	0.979	180	P 1 2_1 1	58.2	139	122	90.2
pAF Nicotinate	Nicotinate	Diamond	I04	0.979	171	P 1 2_1 1	71.5	58.0	117	104

Table 4-2: Data collection statistics for NfsB unnatural amino acid mutant datasets. Diamond refers to Diamond Light Source, UK.

	Processing							
Dataset Identifier	*Resolution Range (Å)*	*No. Observations*	*No. Unique*	*Multiplicity (%)*	*Completeness (%)*	*I/σI*	*R_{sym} (%)*	*Mosaicity*
pNF Acetate	24.32-2.74 (2.91-2.74)	18,5177 (25,790)	49,791 (7,006)	3.72 (3.68)	98.3 (94.5)	10.5 (3.3)	14.6 (56.2)	0.203
pAF Nicotinate	27.72-2.71 (2.88-2.71)	11,3391 (16,384)	32,949 (4,775)	3.44 (3.43)	99.7 (99.7)	11.5 (5.4)	9.4 (26.3)	0.66

Table 4-3: Processing statistics of NfsB unnatural amino acid mutant datasets. Values in parentheses are for the highest resolution shell. $R_{sym} = (\Sigma_{hkl}\Sigma_j|I_{hkl,j} - \langle I_{hkl}\rangle|)/\Sigma_{hkl}\Sigma_j I_{hkl,j}$.

	Refinement						Mean B Factors (Å^2)			
Dataset Identifier	*No. of Reflections*	*No. Water Molecules*	*R factor (%)*	*R_{free} (%)*	*RMSD bond angles (°)*	*RMSD Bond Lengths (Å)*	*Protein*	*Ligand*	*FMN*	*Water*
pNF Acetate	49,768	272	16.42	23.35	1.03	0.007	34.7	66.1	43.1	35.7
pAF Nicotinate	32,937	239	16.23	22.99	0.99	0.007	25.9	28.9	17.5	24.0

Table 4-4: Refinement statistics for NfsB unnatural amino acid mutant datasets. *R-factor* $= (\Sigma|\,|F_{obs}|-|F_{calc}|\,|)/\Sigma|F_{obs}|$.The *R-free* is the R-factors calculated from the 5% of data left out of refinement.

4.8 Structure of pNF mutant bound to acetate

The pNF mutant structure contained eight copies of the NfsB monomer in the ASU. The 124 residue of interest was not as clear in the density as many of the other active site residues. However since the pNF side chain is large and rigid, unambiguous assignment of side chain rotamer was possible. The electron density around the mutation sites in each of the 8 chains is shown in Figure 4-9. In chains F and G the pNF residue was modelled in two alternate conformations, although the dominant rotamer is in the same conformation as all the other monomers.

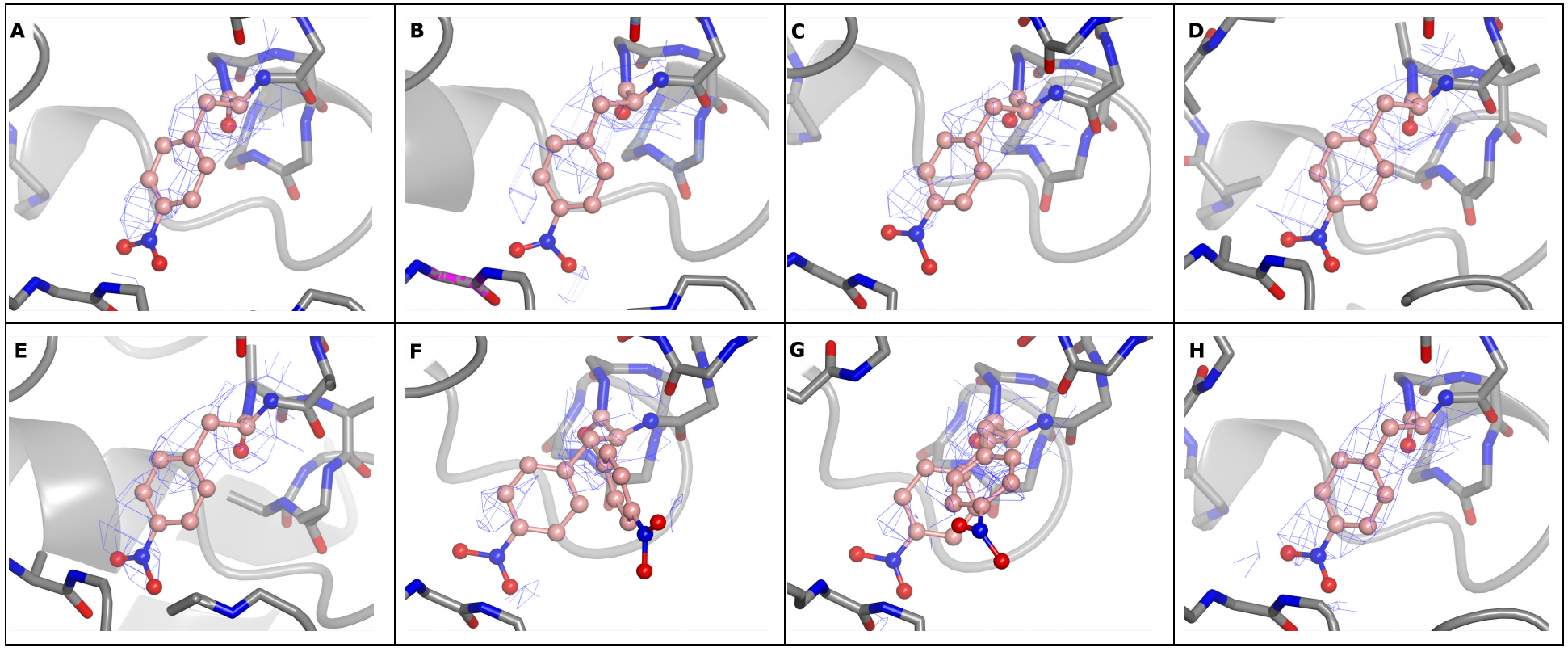


Figure 4-9: Electron density for the mutation sites of pNF NfsB mutant. The pNF124 (pink) is shown for each chain (labelled A-H). The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.0 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

Of the eight active sites, seven contain a bound acetate ligand and the other (site C) contains just water. The electron density for the modelled ligands is shown in Figure 4-10. The conformations of the ligands are difficult to assign unambiguously at this resolution (2.7 Å), but it is clear that there are differences between the positions of the acetate in each active site.

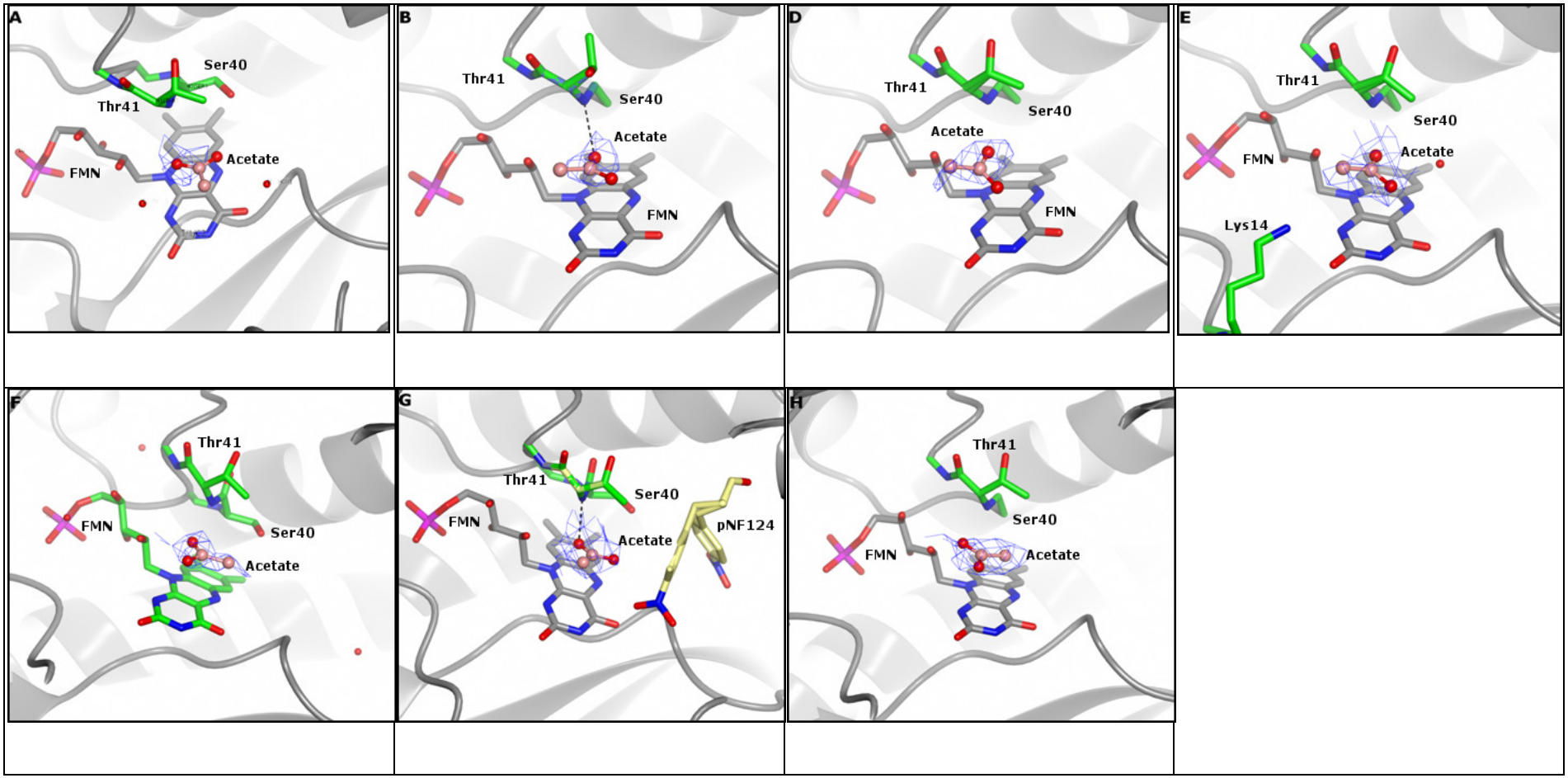


Figure 4-10: Electron density for acetate in the structure of the pNF NfsB mutant. The chain identifier is in the top left hand corner of each image. Active site C was empty and so is not shown. Residues within 5.0 Å of the acetate molecules are shown. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

The mutation of Phe to pNF causes a change in its side chain rotamer. The dominant position of the pNF has changed relative to the position of the Phe ring as shown in Figure 4-11; this is a change from an m-85 rotamer to a t80 rotamer. This shift also causes a 1.0 Å change in the Cα position of Phe124 and less pronounced changes in surrounding residues. The electron density is not sufficient to allow assignment of the exact position of the nitro group of pNF, so the active site H-bonding patterns cannot be described unambiguously. However, comparison of the eight active sites allows us to identify the nature of the interactions that have caused the shift in side chain position. The pNF mutation replaces a hydrophobic residue (Phe) with a more polar residue that is likely negatively charged. The most conserved interaction for pNF is an end-on, π-π stacking interaction between the aromatic ring of the pNF and His128 (~ 3.2 Å). Over each of the seven pNF containing active sites there are other residues that interact with the nitro group; the backbone carbonyl and side chain of Glu165, the side chains of Trp68, Met127, His128 and Trp138 (all shown in Figure 4-11). Of these residues, it would be expected that Trp68, His128, Trp138 would be able to form H-bonds to the oxygens of the nitro group. The movement of the pNF side chain relative to the Phe sidechain also causes the bound acetate to move by around 1.5 Å, away from the pNF sidechain (visible in Figure 4-11).

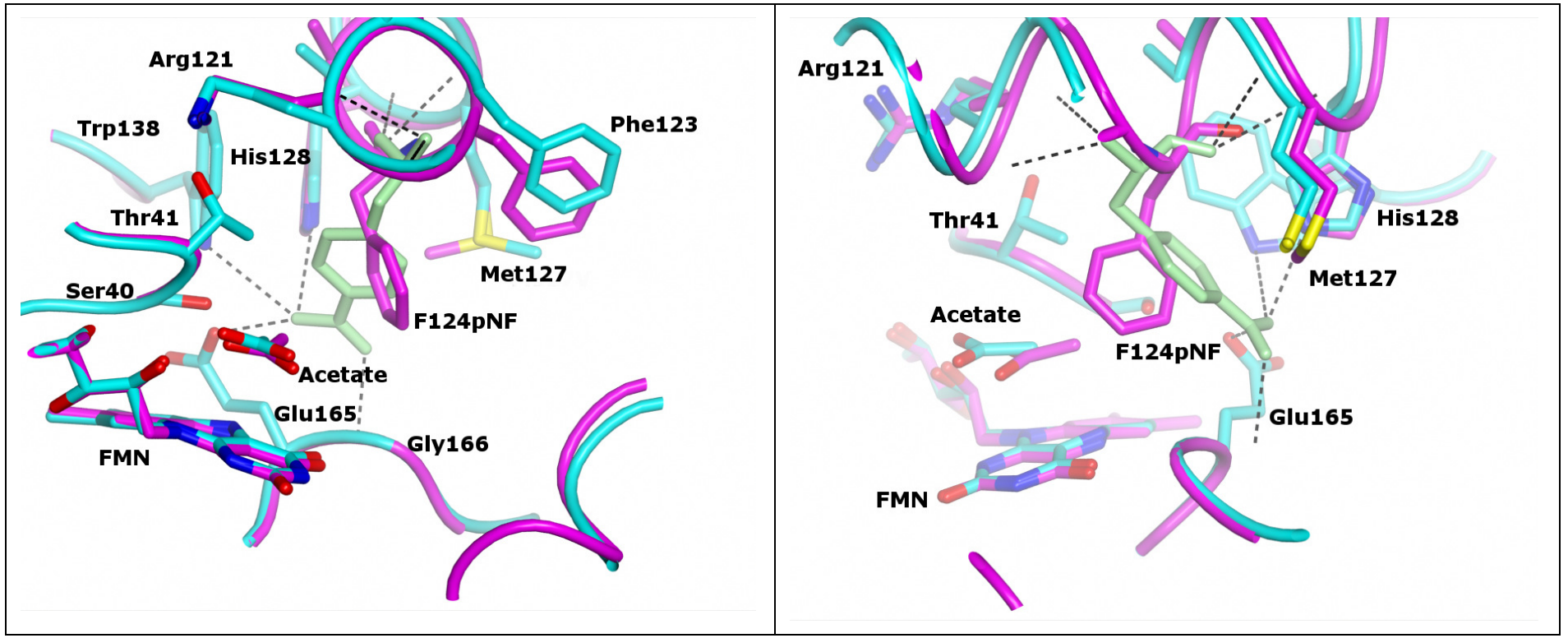


Figure 4-11: Change in the position of the 124 residue when mutated to pNF (light green). WT-NfsB bound to acetate (pdb I.D. 1YLR, (Race *et al.*, 2005)) is shown in purple and the pNF-Acetate structure in light blue. Structures aligned using THESEUS (Theobald and Wuttke, 2008). Images are 90° apart. Figures produced using CCP4 Molecular Graphics.

When the backbone structures are compared to WT-NfsB there are two differences seen. The first of these is a shift in the position of helix F, the helix containing the 124 residue, which is at its most pronounced at the outer edge, where the loop from Arg129 to His133 shifts ~1.5 Å. This can be seen to be caused by the change in effective rotamer caused by the mutation to pNF, forcing the helix away from the adjacent Glu165 and Gly166. There is also a slight change (~1.0 Å) in the position of the Phe70 containing loop, moving it towards the active site and the pNF, as the change in effective rotamer has left space to be filled.

4.9 Structure of pAF mutant bound to nicotinate

The crystal structure of the pAF mutant bound to nicotinate has four copies of the NfsB monomer in the ASU, and thus four crystallographically unrelated active sites. All four copies of the monomer (including the FMN prosthetic group and the nicotinate inhibitor) align well and there are no major differences. The pAF mutations are in clean density comparable with other active site residues and show little variance across the active sites. The electron density of the four mutation sites is shown in Figure 4-12. In all four active sites the pAF has shown the same rotamer change seen with the pNF mutant. The density for the four nicotinate ligands is shown in Figure 4-13, and is clear for each active site.

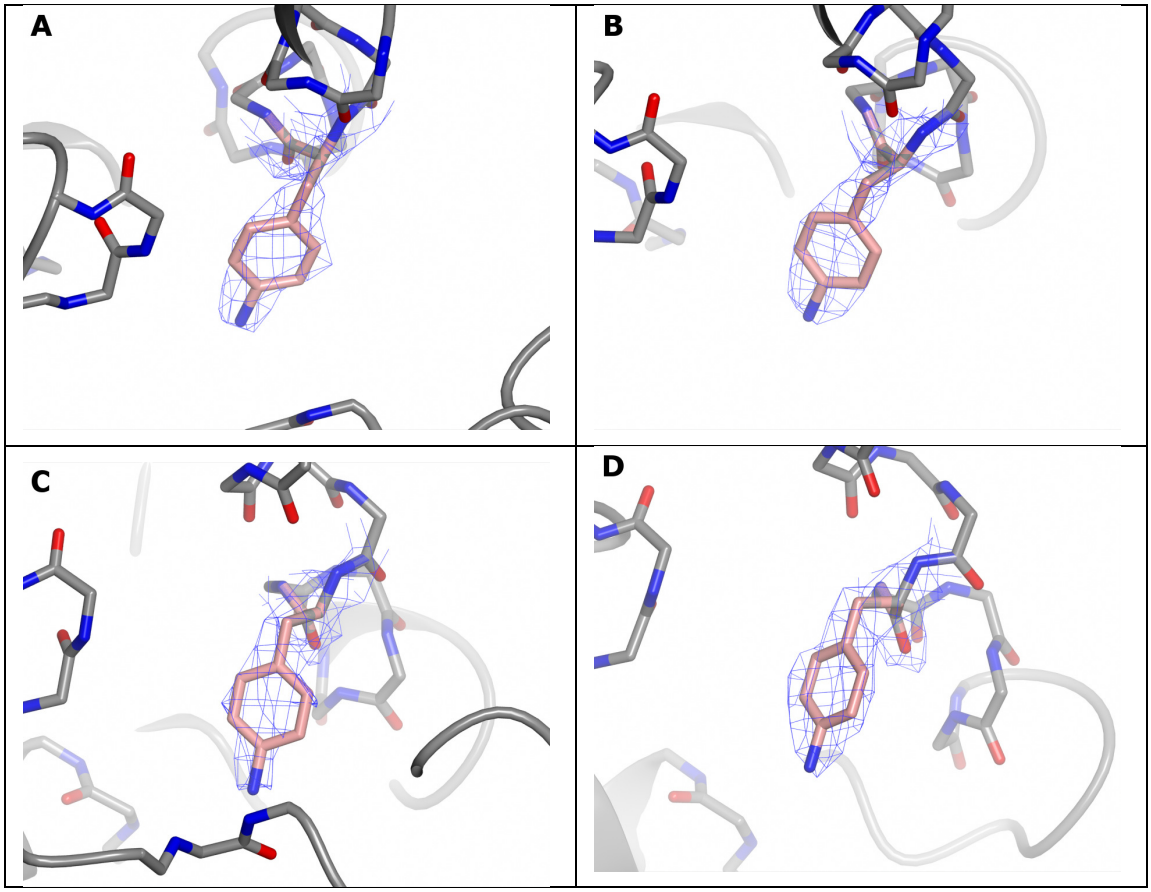


Figure 4-12: Electron density of the pAF (pink) mutation sites in each monomer of the pAF Nicotinate Structure. Images are labelled by chain. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

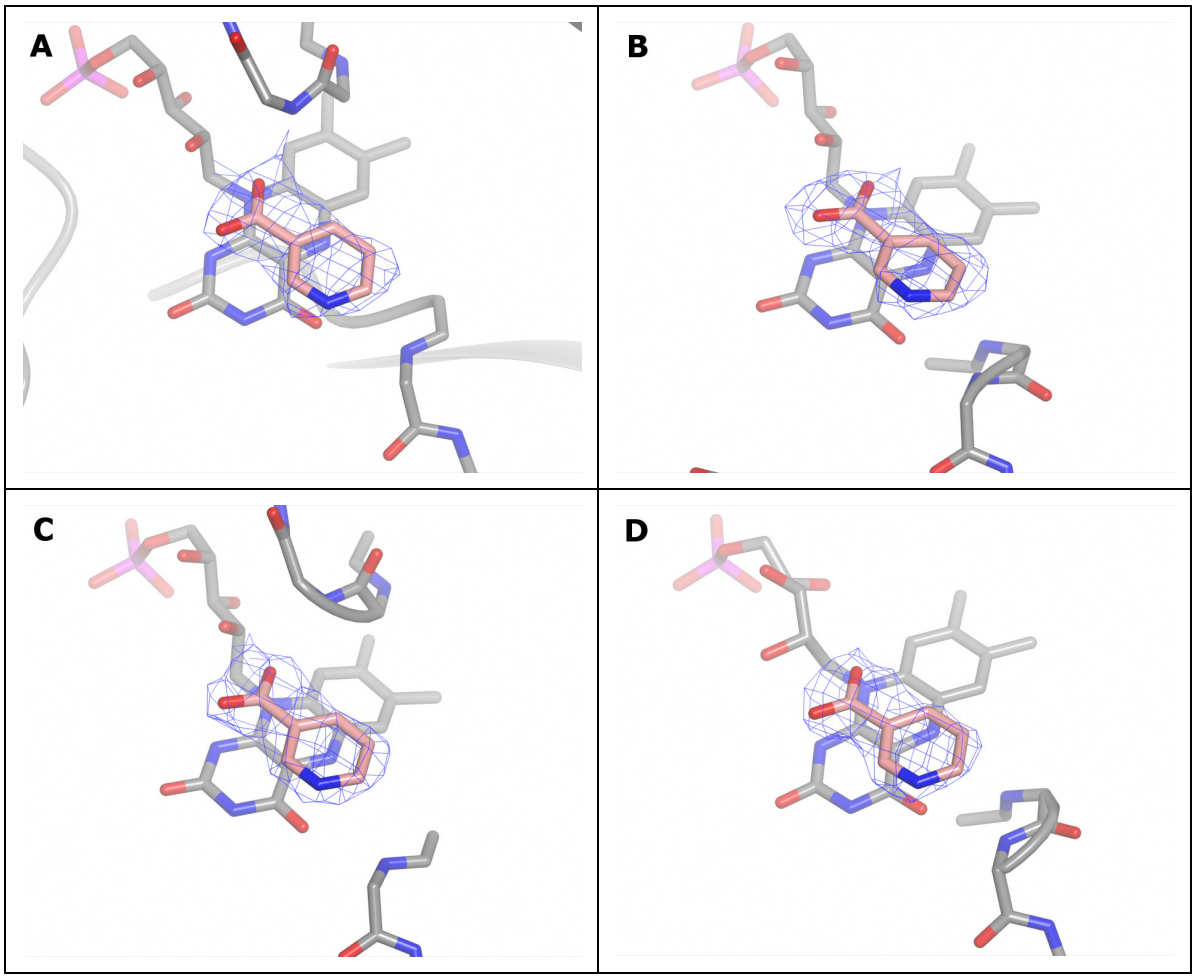


Figure 4-13: Electron density for nicotinate ligands (all in pink) in pAF structure. Images are labelled by chain. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

When the pAF mutant is compared to WT-NfsB there is only one consistent difference, which is a slight shift (< 2.0 Å) in the position of helix F containing the pAF124 residue (residues 111-120) (Figure 4-14). This shift is caused by the change in the side chain rotamer upon mutation to pAF, clearly seen in Figure 4-14, which pushes the helix slightly away from the FMN. The conformational change may be explained by the presence of two H-bonds from the amino group of the pAF to Glu165, one to the backbone carbonyl and one to the side chain carboxylic acid (Figure 4-15). Other residues that interact with the side chain of pAF in this conformation are Met127, His128 and Trp138 (visible in Figure 4-14 and Figure 4-15, which also interact with the pNF mutation (Section 4.8)).

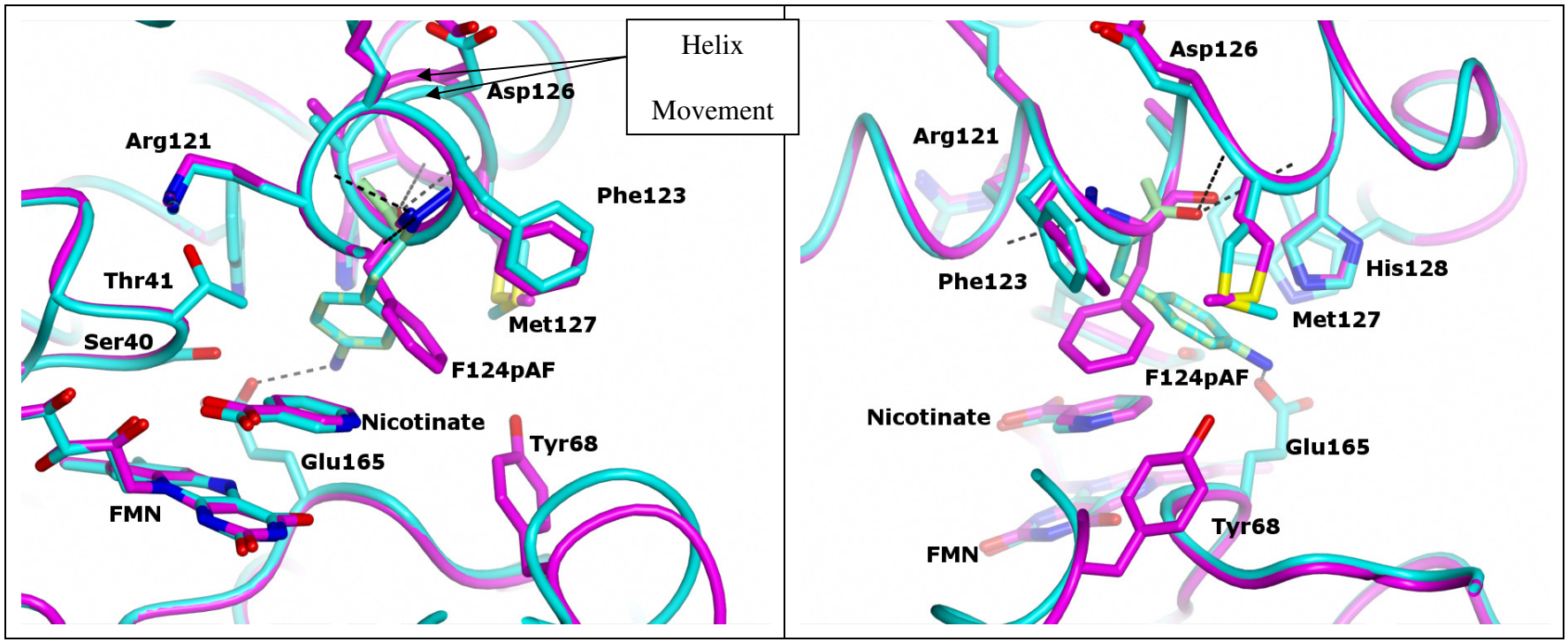


Figure 4-14: Comparison of pAF Nicotinate structure (light blue) with nicotinate bound WT-NfsB (purple - pdb I.D. 1ICR (Lovering *et al.*, 2001)). Structures aligned using THESEUS (Theobald and Wuttke, 2008). The pAF is shown in light green. Two images are 90 ° apart. Figures produced using CCP4 Molecular Graphics.

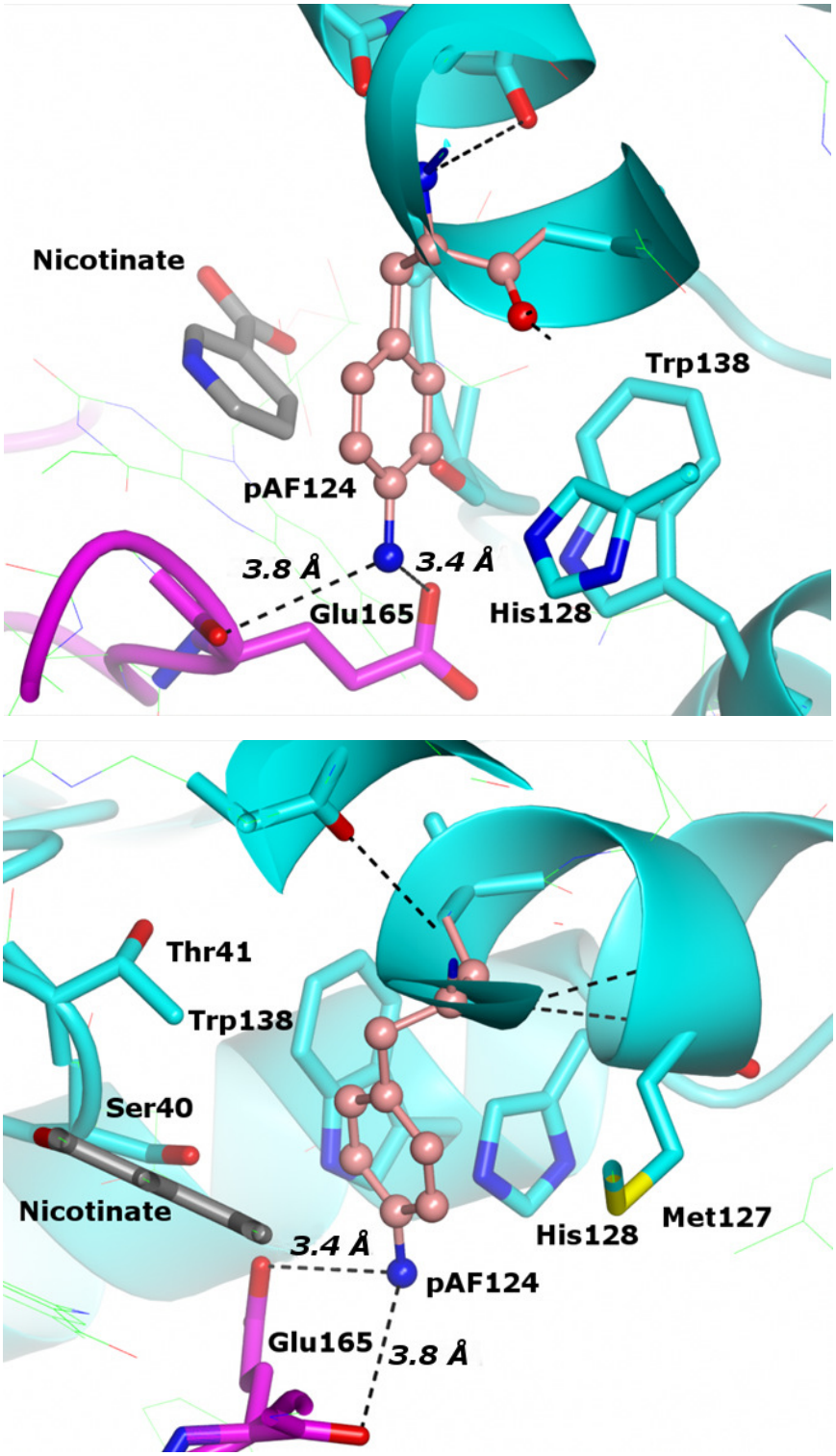


Figure 4-15: Interactions of pAF124 with surrounding residues. Residues shown are within 4.0 Å of pAF124. H-bonds are shown as dashed lines and distances are indicated. Two images separated by 90 °. Figures were produced using CCP4 Molecular Graphics.

4.10 New Crystal Form

Of the various new crystal forms seen with the cleaved His-tagged NfsB, only one, pNF-Acetate, led to a structure that was solved. Table 4-5 shows a comparison of all the currently known crystal forms of NfsB, with the new 'Monoclinic 4' form seen with the pNF mutant with bound acetate. Since the crystals of pNF-acetate are very similar in appearance to those of the pAF-nicotinate, which was in the Monoclinic 1 space group, attempts were made to reprocess the pNF-acetate data in other space groups including those similar to previously seen forms. Molecular replacement failed with all other forms. This new crystal form has eight monomers in the ASU.

Identifier	Space Group	a (Å)	b (Å)	c (Å)	β (°)	Copies in ASU	Example
Tetragonal	P 4_1 2_1 2	55	55	260	-	2	pdb I.D. 1ICR (Lovering *et al.*, 2001)
Orthorhombic 1	P 2_1 2_1 2_1	55	120	145	-	4	pdb I.D. 1ICV (Lovering *et al.*, 2001)
Orthorhombic 2	P 2_1 2_1 2	155	45	55	-	2	pdb I.D. 1YLU (Race *et al.*, 2005)
Monoclinic 1	P 1 2_1 1	70	55	115	105	4	pdb I.D. 1ICU (Lovering *et al.*, 2001) pAF Nicotinate
Monoclinic 2	P 1 2_1 1	55	60	75	100	2	Triple SN27683
Monoclinic 3	C 1 2 1	115	45	95	115	2	Double SN27683
New Forms:							
Monoclinic 4	P 1 2_1 1	58	139	122	90.2	8	pNF Acetate

Table 4-5: Summary of the crystal forms seen with NfsB, natural amino acid mutants and unnatural amino acid mutants. For previously seen forms examples are given from the literature, for the new forms examples are given from this study. Average cell dimensions are given to the nearest 5 Å as there is up to 5 Å difference between crystals of the same form, other than for the new monoclinic form since there is only one current structure.

4.11 Steady-State Kinetics

Steady-state kinetics have been measured for various substrates of WT-NfsB and its natural mutants (Race *et al.*, 2005, Race *et al.*, 2007, Jarrom *et al.*, 2009). For the unnatural mutants, only experiments comparing the specificity constant (k_{cat}/K_m) for CB1954 have been reported (Jackson *et al.*, 2006). Here we determine the k_{cat}/K_m for

CB1954 and Menadione, for these proteins, both at two NADH concentrations. In order to see whether the His-tag or cleavage of the His-tag has any effect on the activity, three stocks of WT enzyme were tested: - first one expressed with a His-tag and purified as described in Section 2.3, second WT-NfsB with His-tag cleaved as described in Section 2.3.3, and finally native WT-NfsB expressed without a His-tag as described in Section 2.2. Finally the steady-state kinetics of the pAF and pNF mutants have been studied. Both of these preparations had the His-tag cleaved off and comparisons with WT are with the cleaved form. All cleaved samples showed single band purity on SDS-PAGE and were in the presence of enterokinase for the same time, however any His-tagged protein was not removed using Nickel NTA. This would ensure that only the cleaved protein remained and should ideally be done before any future experiments. Experiments were performed for WT NfsB by Asha Hassan (MRes project student), for the pNF mutant by Sophie Englander and for the pAF mutant by Laura Goodman (both undergraduate project students). All under the authors direct supervision.

4.11.1 CB1954 Kinetics

When analysing the steady-state kinetics of NfsB and CB1954, it is often not possible to get a full Michaelis-Menton curve due to the low solubility of CB1954 in solution and its high K_m. The later part of the curve is often not seen at all for WT NfsB, but as the K_m of the mutants increases, more of the curve is seen. If the K_m is very high it is not possible to fit the data to Equation 2-1 accurately to determine k_{cat} and K_m separately. However the data can still be fitted to Equation 2-4, and k_{cat}/K_m (the specificity constant) can be determined accurately as this is the tangent to the initial

slope of the curve. The k_{cat}/K_m is equivalent to the initial gradient of the Michaelis-Menten curve, and for an enzyme displaying ping-pong kinetics it should be independent of NADH concentration. So comparing the k_{cat}/K_m gives an indication of the relative binding strengths.

4.11.1.1 Effect of His-tag on the Kinetics of WT NfsB

The reduction of CB1954 by the untagged (native), His-tagged and cleaved WT-NfsB at 60 µM NADH is shown in Figure 4-16 and at 200 µM NADH in Figure 4-17. The statistics of the fit to the hyperbola are summarised in Table 4-6. The curve fits are good with the exception of the untagged protein with 60 µM NADH. The rates are all higher at the higher NADH concentration as would be expected. When comparing the different WT proteins, there is not a large difference between any of them. However the His-tagged protein does have the lowest k_{cat}/K_m of the proteins, implying there is a slight, negative effect of the His-tag on the activity of the enzyme which is restored upon cleavage, but this is within the error of the measurements.

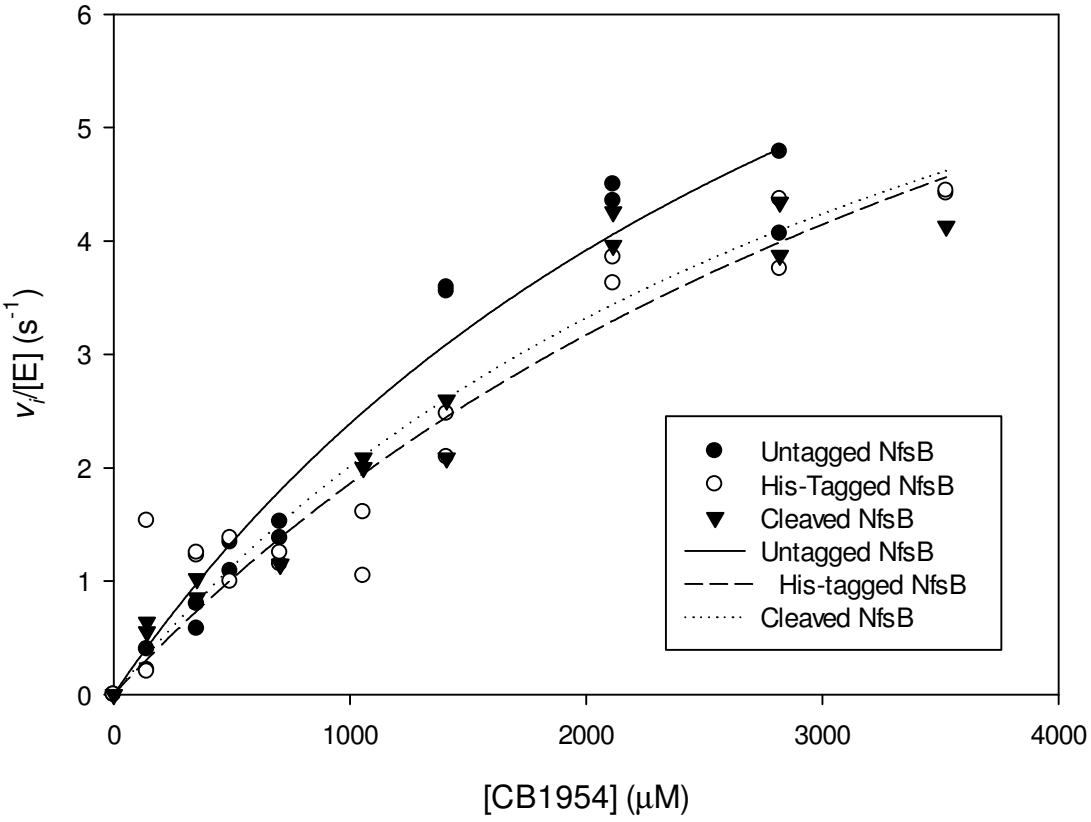


Figure 4-16: Comparison of the steady-state reduction of CB1954 by WT-NfsB at 60 μM NADH, in untagged, His-tagged and cleaved forms.

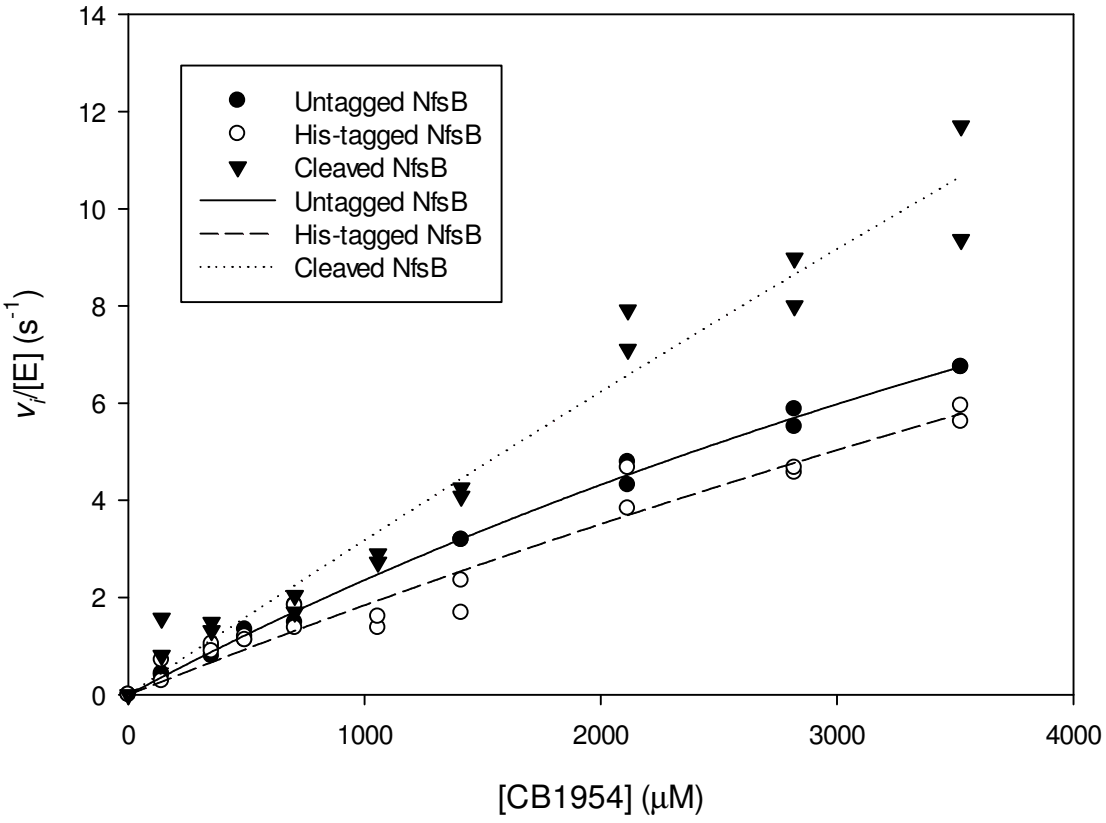


Figure 4-17: Comparison of the steady-state reduction of CB1954 by WT-NfsB at 200 μM NADH, in untagged, His-tagged and cleaved forms.

Protein	[CB1954] µM	[NADH] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
Untagged NfsB	141-2820	60	10.9 ± 2.7	4.0	0.0017	3560 ± 1370	2.6	0.023	0.0031 ± 0.0004	7.0	<0.0001
Untagged NfsB	141-3525	200	25.7 ± 3.5	7.4	<0.0001	9910 ± 1710	5.8	<0.0001	0.0026 ± 0.0001	26	<0.0001
His-tagged NfsB	141-3525	60	10.7 ± 3.4	3.1	0.0064	4780 ± 2270	2.1	0.052	0.0022 ± 0.0004	6.0	<0.0001
His-tagged NfsB	141-3525	200	ND	-	-	ND	-	-	0.0019 ± 0.0002	8.4	<0.0001
Cleaved NfsB	141-3525	60	9.5 ± 2.3	4.2	0.0011	3710 ± 1410	2.6	0.020	0.0026 ± 0.0004	6.8	<0.0001
Cleaved NfsB	141-3525	200	ND	-	-	ND	-	-	0.0032 ± 0.0004	8.5	<0.0001

Table 4-6: Steady-state kinetic parameters for the reduction of CB1954 by WT-NfsB (untagged, His-tagged and cleaved) with NADH. Each set of experiments were performed at a constant NADH concentration and at variable CB1954 concentration. ND – Not determined; data cannot be fitted accurately to a simple hyperbola.

4.11.1.2 Steady-State Kinetics of pNF and pAF mutants with CB1954

The cleaved pNF and pAF NfsB mutants were both assayed for the reduction of CB1954 at 60 µM and 200 µM. A comparison of the two mutants and the cleaved WT NfsB at 60 µM is shown in Figure 4-18, and at 200 µM in Figure 4-19. Apparent Michaelis-Menten parameters for pNF and pAF are detailed in Table 4-7 and Table 4-8 respectively. Both mutants show improvement over WT as expected from previous reports (Jackson *et al.*, 2006), although the exact k_{cat}/K_m values (consistent across the two NADH concentrations) are about 2-fold lower than reported Jackson *et al.*. As previously seen both mutants appear to be about 20-fold more active than WT, with the pNF showing the greater improvement in specificity constant for CB1954.

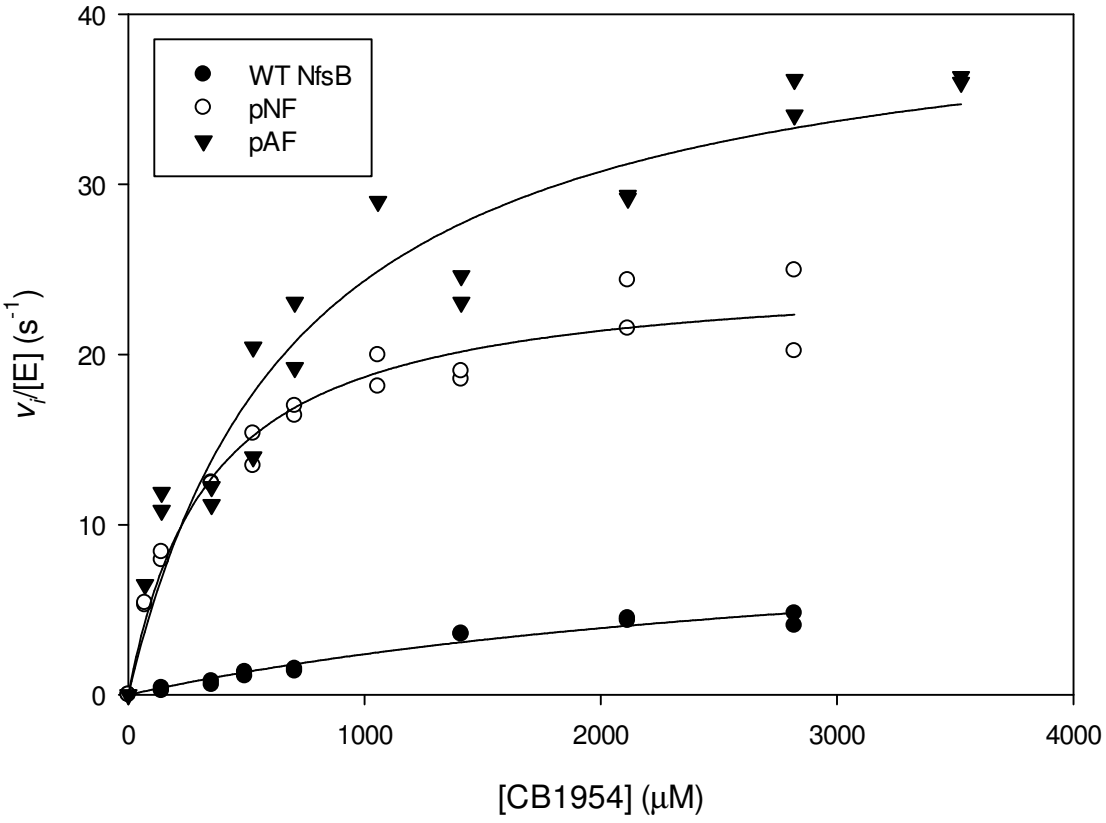


Figure 4-18: Reduction of CB1954 by cleaved WT NfsB and pNF and pAF mutants. All experiments performed at 60 μM NADH.

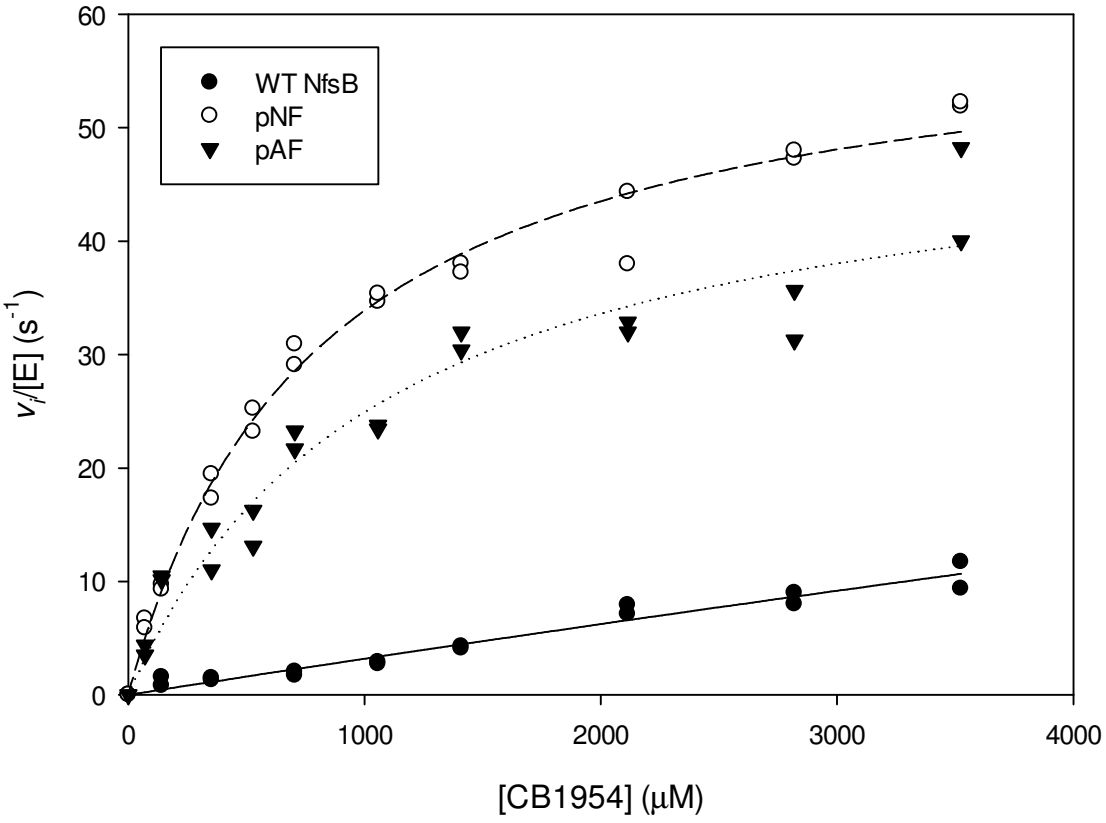


Figure 4-19: Reduction of CB1954 by cleaved WT NfsB and pNF and pAF mutants. All experiments performed at 200 μM NADH.

[NADH] µM	[CB1954] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
60	141-2820	25.0 ± 1.0	25	<0.0001	341 ± 48	7.1	<0.0001	0.073 ± 0.0080	9.2	<0.0001
200	141-3525	60.9 ± 1.9	32	<0.0001	797 ± 70	11	<0.0001	0.076 ± 0.0046	17	<0.0001

Table 4-7: Kinetic parameters for the reduction of CB1954 by the pNF NfsB mutant.

[NADH] µM	[CB1954] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
60	141-3525	41.8 ± 2.8	15	<0.0001	716 ± 140	5.0	<0.0001	0.058 ± 0.0083	7.0	<0.0001
200	141-3525	51.6 ± 4.2	12	<0.0001	1070 ± 220	5.0	0.0001	0.048 ± 0.0062	7.7	<0.0001

Table 4-8: Kinetic parameters for the reduction of CB1954 by the pAF NfsB mutant.

4.11.2 Menadione Kinetics

4.11.2.1 Effect of His-tag on the Steady-State Kinetics of WT NfsB

The three WT proteins (untagged, His-tagged and cleaved) were tested with menadione as a substrate, as they were with CB1954 (Section 4.11.1.1). The steady-state kinetics results are shown at 60 μM NADH in Figure 4-20 and at 200 μM NADH in Figure 4-21. The statistics of the fit to the Michaelis-Menten equation are shown in Table 4-9. The specificity constants show again that the His-tagged NfsB has less activity than untagged protein, but cleavage of the His-tag does not appear to restore the full activity with menadione, unlike with CB1954 as a substrate. This may be because menadione is a better substrate so small differences are more noticeable. The values for the cleaved protein at 60 μM NADH are much lower than would be expected from the other results, mostly in terms of k_{cat}.

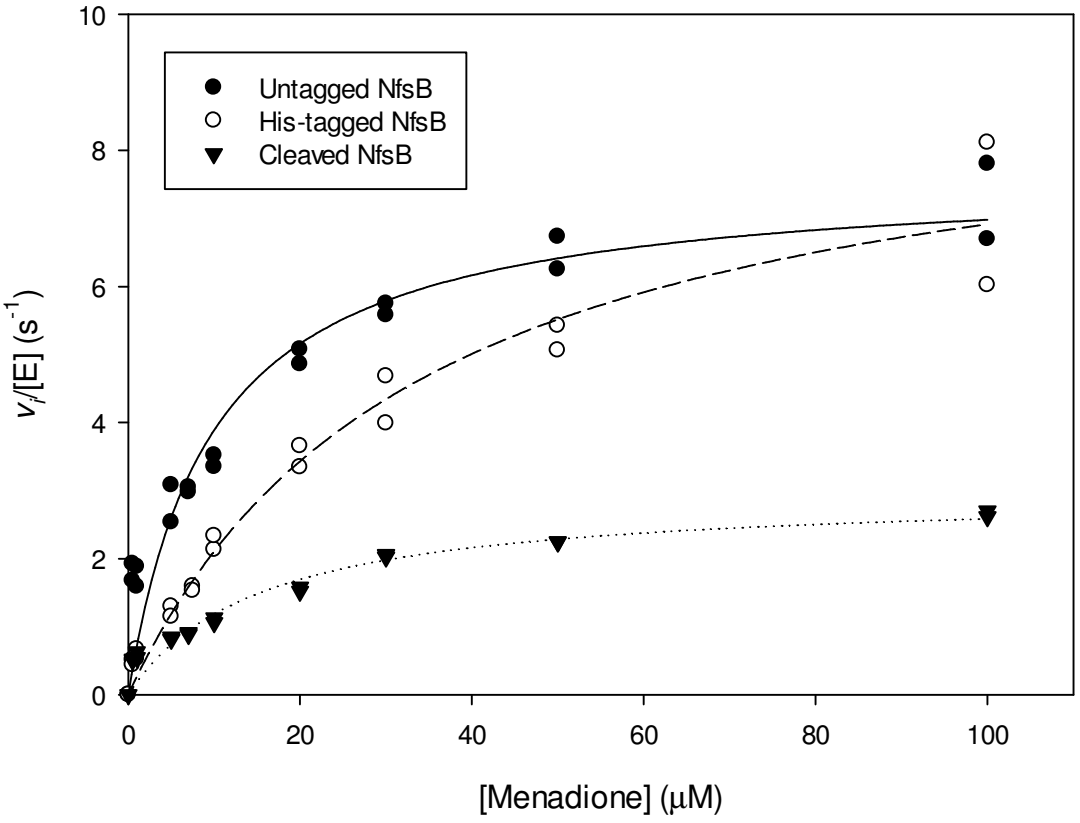


Figure 4-20: Comparison of the steady-state reduction of menadione by WT-NfsB, in untagged, His-tagged and cleaved forms at 60 μM NADH.

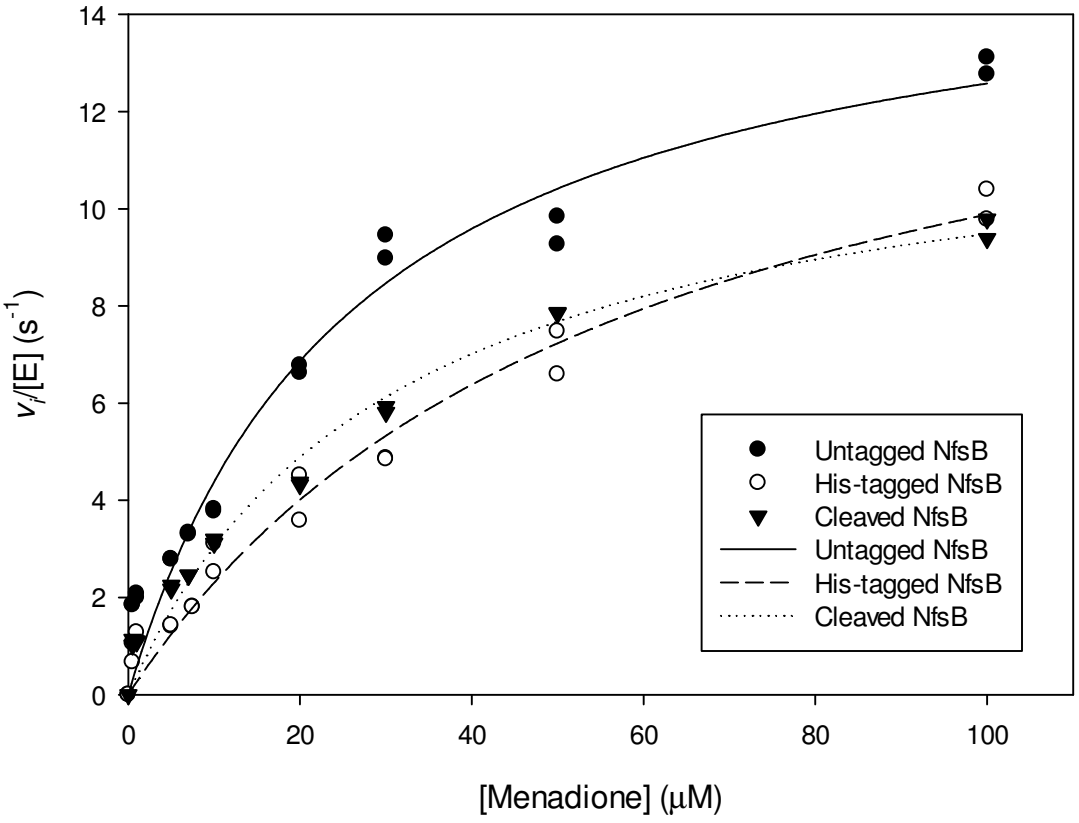


Figure 4-21: Comparison of the steady-state reduction of menadione by WT-NfsB, in untagged, His-tagged and cleaved forms at 200 μM NADH.

Protein	[Menadione] µM	[NADH] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
Untagged NfsB	0.5 - 100	60	7.65 ± 0.51	15	<0.0001	9.66 ± 2.1	4.6	0.0003	0.79 ± 0 .13	6.0	<0.0001
Untagged NfsB	0.5 - 100	200	15.9 ± 1.2	13	<0.0001	26.2 ± 4.7	5.6	<0.0001	0.61 ± 0.071	8.6	<0.0001
His-tagged NfsB	0.5 - 100	60	9.26 ± 0.73	13	<0.0001	34.0 ± 5.9	5.7	<0.0001	0.27 ± 0.029	9.4	<0.0001
His-tagged NfsB	0.5 - 100	200	15.6 ± 1.5	10	<0.0001	57.6 ± 11	5.4	<0.0001	0.27 ± 0.026	10	<0.0001
Cleaved NfsB	0.5 - 100	60	2.97 ± 0.20	15	<0.0001	15.1 ± 2.9	5.1	<0.0001	0.20 ± 0.028	15	<0.0001
Cleaved NfsB	0.5 - 100	200	12.4 ± 0.73	17	<0.0001	30.6 ± 4.2	7.3	<0.0001	0.40 ± 0.035	12	<0.0001

Table 4-9: Steady-state kinetic parameters for the reduction of menadione by WT-NfsB (untagged, His-tagged and cleaved) with NADH. Each set of experiments were performed at a constant NADH concentration and at variable menadione concentration.

4.11.2.2 Mutant Kinetics

The reduction of menadione by the pNF and pAF mutants was also measured and compared to WT at 60 µM (Figure 4-22) and 200 µM NADH (Figure 4-23). The steady-state parameters are detailed in Table 4-10 and Table 4-11 respectively. Both displayed substrate inhibition above 25 µM and 50 µM respectively, and so data points beyond this were not used for fitting. This has the effect that the full curve is not seen at 200 µM NADH for the pNF mutant, making the estimates of k_{catapp} and K_{mapp} in particular poorer than other experiments. However the P values are still below 0.0001 and the t values above 5. The improvement in k_{cat}/K_m seen for the pAF mutant over WT is greater than those seen with CB1954, but the pNF mutant shows a similar degree of improvement (~ 20-fold). The most improved mutant for menadione is the pAF, rather than the pNF, with ~ 50-fold improvement over WT.

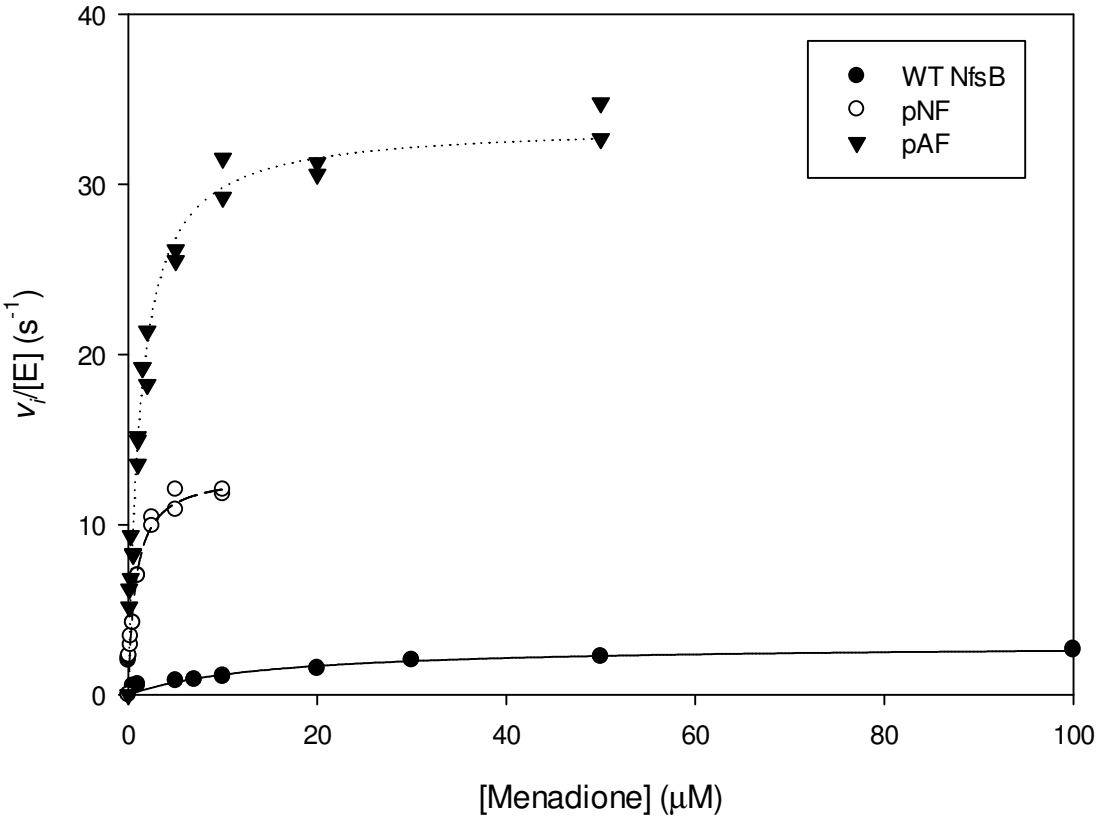


Figure 4-22: Reduction of menadione by cleaved His-tagged WT NfsB and pNF and pAF mutants. All experiments performed at 60 μM NADH.

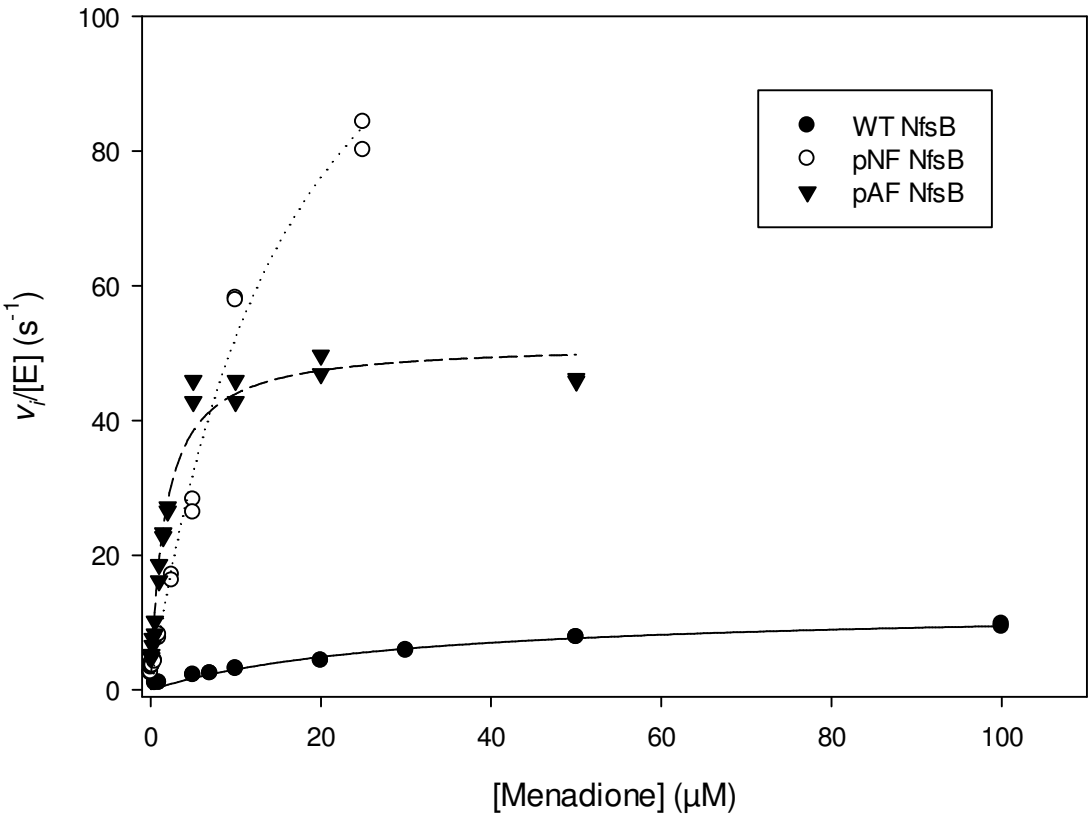


Figure 4-23: Reduction of menadione by cleaved His-tagged WT NfsB and pNF and pAF mutants. All experiments performed at 200 μM NADH.

[NADH] µM	[Menadione] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
60	0.01- 10	13.0 ± 0.67	20	<0.0001	0.789 ± 0.14	5.5	<0.0001	16.5 ± 2.4	6.8	<0.0001
200	0.01 - 25	139 ± 10	13	<0.0001	16.6 ± 2.3	7.1	<0.0001	8.39 ± 0.61	14	<0.0001

Table 4-10: Kinetic parameters for the reduction of menadione by the pNF NfsB mutant.

[NADH] µM	[Menadione] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
60	0.1 – 50	33.5 ± 0.85	39	<0.0001	1.24 ± 0.12	11	<0.0001	27.1 ± 2.1	13	<0.0001
200	0.1 – 50	51.5 ± 1.5	34	<0.0001	1.71 ± 0.18	9.3	<0.0001	30.2 ± 2.7	11	<0.0001

Table 4-11: Kinetic parameters for the reduction of menadione by the pAF NfsB mutant.

4.11.3 Comparison of Kinetics with Previous Mutants

Table 4-12 compares the enzyme fractions presented in this chapter with previously published data for WT and the most active single, double and triple mutants. There is variation in the WT NfsB results between this study and previously published data, but each of the previous mutants has a specificity constant k_{cat}/K_m. for CB1954 at least an order of magnitude higher than for the WT NfsB in this study. The two unnatural amino acid mutants studied here show a greater improvement with CB1954 than the best natural single mutant (F124N), but are not as good as the best double and triple mutants (T41L/N71S and T41Q/N71S/F124T). With menadione, the pAF mutant is the most active of those detailed, with the pNF mutant also showing greater improvement than with CB1954. Thus the selectivity for CB1954 is worse for pAF than for WT protein, although they do bind CB1954 more tightly. The pNF mutant shows a similar degree of selectivity to WT NfsB and F124N.

Protein	$k_{cat}/K_{m\text{CB1954}}$ ($\mu M^{-1} s^{-1}$)	$k_{cat}/K_{m\text{menadione}}$ ($\mu M^{-1} s^{-1}$)	Selectivity for CB1954	Reference
WT NfsB (Untagged)	0.0026 ± 0.0001	0.61 ± 0.071	0.0043	This Study – (200 µM NADH)
WT NfsB (His-Tagged)	0.0019 ± 0.0002	0.27 ± 0.026	0.0070	This Study – (200 µM NADH)
WT NfsB (Cleaved)	0.0032 ± 0.0004	0.40 ± 0.035	0.0080	This Study – (200 µM NADH)
WT NfsB (Previous)	0.007 ± 0.0006	0.69 ± 0.05	0.010	(Race *et al.*, 2005)
F124N	0.031 ± 0.006	3.0 ± 0.6	0.010	(Jarrom *et al.*, 2009)
T41L/N71S	0.71 ± 0.08	21.7 ± 3.7	0.033	(Jarrom *et al.*, 2009)
T41Q/N71S/F124T	0.318 ± 0.016	18.8 ± 3	0.017	(Jarrom *et al.*, 2009)
pNF	0.076 ± 0.0046	8.39 ± 0.61	0.0091	This Study – (200 µM NADH)
pAF	0.048 ± 0.0062	30.2 ± 2.7	0.0016	This Study – (200 µM NADH)

Table 4-12: Comparison of specificity constants (k_{cat}/K_m) of the WT NfsB used in this study, previous WT NfsB, the most improved single double and triple mutants, and the unnatural amino acid mutants, pNF and pAF. Selectivity is the values for CB1954 divided by the values for menadione.

4.12 Conclusions

In Section 1.6.2.3.4 the use of unnatural amino acids to improve the activity of NfsB with CB1954 was discussed. In this chapter the structure of the two most improved unnatural amino acid mutants, replacing Phe 124 with pNF and pAF, was shown. The structures show how an effective change in side chain rotomer allows the new chemical groups to interact with the flavin and ligands. The pNF mutant is the most improved single mutant to date for CB1954, whilst pAF is the most improved for menadione. Both mutants show greater improvements with menadione than with CB1954. In both unnatural amino acid mutants there is an effective change of side chain rotamer compared to WT; this change places the new polar group adjacent to the active site where it could interact with a substrate. The change is almost the opposite of the changes caused by previous Phe124 mutations, F124K and F124N, moving towards the FMN prosthetic group rather than away. The rotamer change also causes slight shifts which open the mouth of the active site. The pAF side chain modelled much more clearly and with the relative improvements seen with CB1954 and menadione, it would appear that this change in conformation is more significant to the reduction of the larger menadione, perhaps providing a larger area for the substrate to enter the active site. These new chemical groups may affect catalysis by interacting with the FMN, creating more space in the active site or by providing a group to H-bond to the substrate.

To date (07/10/12) there are fewer than 50 crystal structures in the protein data bank incorporating unnatural amino acids and many of these are duplicates. These few include T4 lysozyme with p-acetyl-*L*-phenylalanine, pdb I.D.3HWL (Fleissner *et al.*,

2009). There is one structure containing pAF, an Anti-HIV antibody, pdb I.D. 3IDN (Bryson *et al.*, 2009). In this structure, the pAF replaces a Lys and changes to mimic the interactions of a His residue from another chain, which was used as evidence for the importance of conservation of function in that region of the protein. At the time of writing there is no structure in the protein data bank containing pNF.

Unnatural amino acids present opportunities for scientists to investigate the chemistry within the protein environment. Substitution of amino acids with synthetic equivalents expands the range of chemical and steric groups that can be inserted (beyond the 20 natural amino acids). These groups can introduce new chemistry into the active site, and aid our understanding of enzyme mechanisms. However, these unnatural amino acid mutants would be unfeasible for use in GDEPT with CB1954 (Section 1.4). It would involve the addition of extra genes, allowing their incorporation, and would require delivery of the synthetic amino acid, creating new challenges in delivery. The cost of such a treatment would also be high, as the synthetic amino acids are expensive. So whilst unnatural amino acids offer opportunities to gain insights into the chemistry of a protein, their use in a clinical environment is unlikely.

5. NfsA Structures and Kinetics

5.1 NfsA as a Prodrug-Activating Enzyme

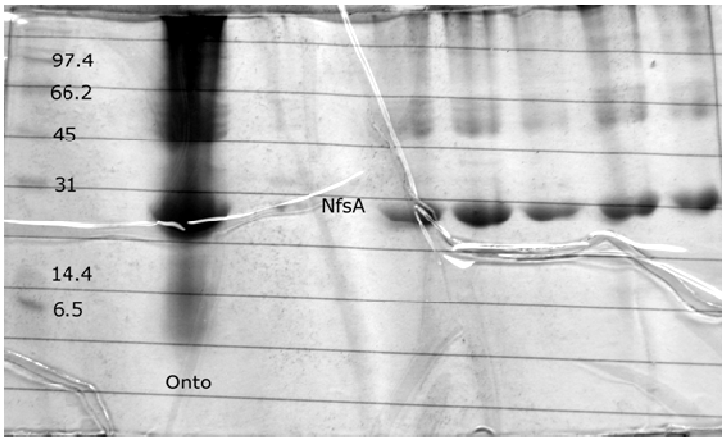
The major nitroreductase of *E. coli*, NfsA, has recently been proposed as an alternative enzyme for use in GDEPT with CB1954 (Vass *et al.*, 2009). The crystal structure of NfsA has previously been determined to 1.7 Å resolution with no ligand modelled in the active site (Kobori *et al.*, 2001). Previous work in our group has elucidated the structure in new conditions, but an unknown ligand was found in the active site (Jarrom, 2008). The density most closely resembled the dicarboxylic acids succinate and fumarate, both later found to be inhibitors of NfsA, but neither was included in the crystallization solution and the density was not clear enough to be unambiguous. Addition of succinate was found to improve the quality of the crystals. In this chapter we present the crystal structures of NfsA without ligands and bound to a number of ligands, including inhibitors and two substrates, further aiding our understanding of the residues involved in the active site and giving clues to mechanism of reaction.

5.2 Purity of purified samples

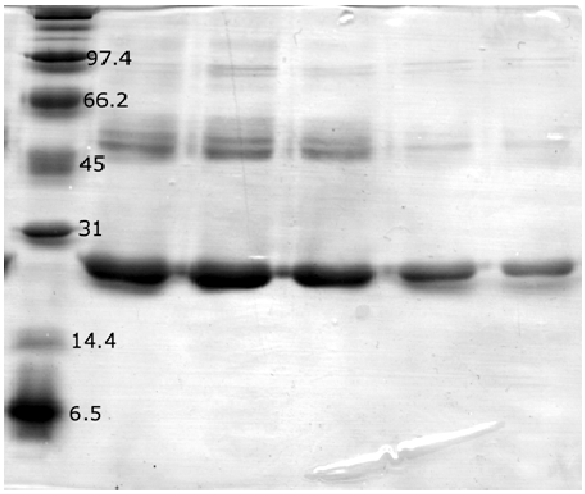
The NfsA protein was expressed in *E. coli* and purified as described in Section 2.4, using ammonium sulphate precipitation followed by phenyl-sepharose and Q-sepharose chromatography, as described previously (Vass *et al.*, 2009). If further purification was required a Blue Trisacryl affinity chromatography column was used, although protein of higher purity failed to crystallize. SDS-PAGE gels of fractions

after each column are shown in Figure 5-1. Attempts to purify further with hydroxyapatite, which had previously been used with NfsB, were unsuccessful.

a)



b)



c)

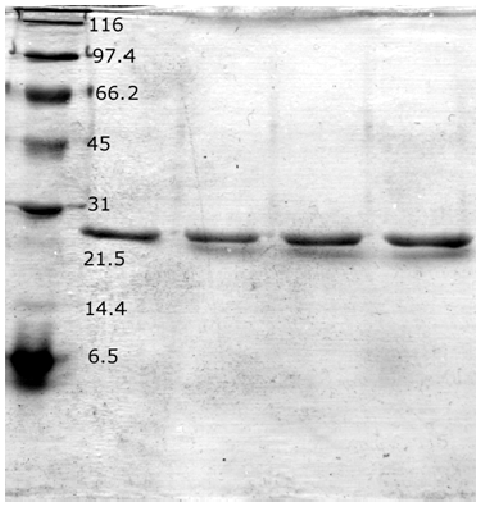


Figure 5-1: SDS-PAGE gels of the purification of NfsA (the brightest band on each), after purification with a) Phenyl-Sepharose b) Q-Sepharose and c) Blue-Tris Acryl matrices. The left hand lane shows the molecular marker with the weights of visible bands marked in kDa. a) Also shows an unpurified fraction, labelled as 'Onto'. The other lanes contain fractions eluted from each column.

5.3 Crystal growth

Commercial crystallization screens and screens based around previously published conditions (Kobori *et al.*, 1999) produced no crystal growth. However conditions based upon those used previously in this laboratory (Jarrom, 2008) led to crystal formation. Crystals appeared within 24 hours and generally reached full size within 48 hours. Typical NfsA crystals are shown in Figure 5-2. All crystals were grown by a sitting-drop method. In general, larger drops produced fewer crystals and less precipitate so the most common successful drop size was 1 µl of protein solution mixed with 1 µl of mother liquor.

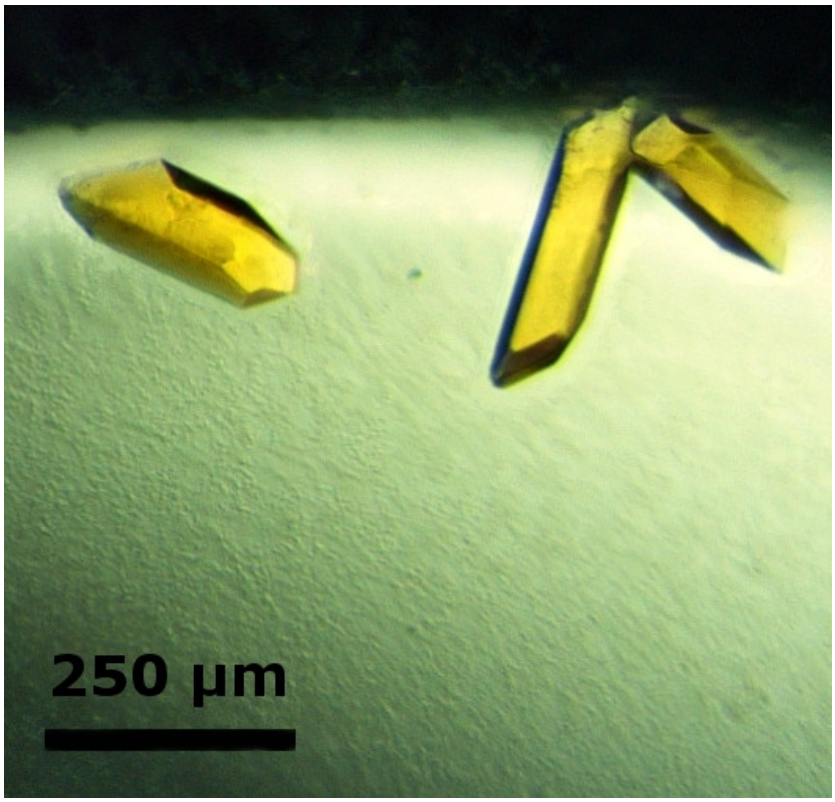


Figure 5-2: An example of crystals of NfsA.

Crystals grew in the presence of up to 10 mM CB1954 (along with 3 % NMP and 11 % PEG 300 used as solvent) and up to 10 mM of the other dinitrobenzamide prodrugs shown in Figure 2-3 (along with 10 % acetonitrile used as a solvent). However the electron density seen in the active site of the crystal structures was ambiguous and didn't improve upon refinement and so these structures were not fully refined.

5.4 Summary of datasets

The crystallization conditions and soaking details for the crystals that give rise to the structures detailed in this chapter are shown in Table 5-1. The data collection statistics for NfsA crystal structures are shown in Table 5-2, the data processing statistics in Table 5-3 and the refinement statistics in Table 5-4. Chemical structures of the ligands found in these structures are shown in Figure 5-3.

	Crystallization – Mother Liquor					Soak	
Dataset Identifier	*Precipitant*	*Imidazole Buffer pH 7.0 (mM)*	*Additive 1*	*Additive 2*	*Protein Conc. (mg/ml)*	*Cryo. Prot.*	*Ligand*
NfsA Succinate 1	30% PEG 3,350	100	50 mM Succinate	3 mM $K_2Fe(CN)_6$	12	15 % Ethylene Glycol	n/a
NfsA Succinate 2	30% PEG 3,350	100	50 mM Malonate	n/a	10	15 % Ethylene Glycol	n/a
NfsA Fumarate	26% PEG 3,000	100	30 mM Fumarate	n/a	11	15 % Ethylene Glycol	n/a
NfsA Empty	26% PEG 3,000	50	n/a	n/a	11	15 % Ethylene Glycol	n/a
NfsA Ethylene Glycol	30% PEG 3,350	100	8.5 mM CB1954*	3 mM $K_2Fe(CN)_6$	10	15 % Ethylene Glycol	n/a
NfsA FMN	30% PEG 3,000	100	20 mM Hydroquinone		10	15 % Ethylene Glycol	n/a
NfsA Quinol	28% PEG 3,000	100	n/a	n/a	9	20 % DMSO	25 mM Hydroquinone
NfsA Quinone	18% PEG 3,000	100	20% DMSO	n/a	7	n/a	10 mM Benzoquinone
NfsA Nitrofurantoin	22% PEG 3,000	100	20 % DMSO	3.75 mM Nitrofuran-toin	11	n/a	n/a

Table 5-1: Crystallization conditions for crystals giving rise to NfsA datasets. The active site contents are indicated in the dataset identifier. All buffers and additives were at pH 7.0. *For the NfsA Ethylene Glycol structure there was also 2.8 % NMP and 10 % PEG 300 used to solubilise the CB1954.

	General			Unit Cell				
Dataset Identifier	*X-ray Source*	*Wavelength (Å)*	*Degrees of Data Collected (°)*	*Space Group*	*a (Å)*	*b (Å)*	*c (Å)*	*β (°)*
NfsA Succinate 1	Home Source	1.54	360	C 1 2 1	91.6	51.6	64.7	134
NfsA Succinate 2	ESRF ID14-1	0.934	360	C 1 2 1	92.2	51.8	64.8	134
NfsA Fumarate	Home Source	1.54	360	C 1 2 1	92.1	52.2	65.0	134
NfsA Empty	ESRF ID23-1	1.00	160	C 1 2 1	91.9	52.1	64.8	134
NfsA Ethylene Glycol	Home Source	1.54	360	C 1 2 1	91.3	51.9	64.4	134
NfsA FMN	ESRF ID14-1	0.934	360	C 1 2 1	91.9	52.0	64.7	134
NfsA Quinol	Home Source	1.54	220	C 1 2 1	92.0	52.0	64.8	134
NfsA Quinone	Home Source	1.54	360	C 1 2 1	91.8	51.9	64.7	134
NfsA Nitrofurantoin	ESRF ID23-1	1.00	180	C 1 2 1	91.4	51.9	64.6	134

Table 5-2: Data collection statistics for NfsA datasets. The active site contents (according to what has been modelled in) are indicated in the dataset identifier. Home Source, refers to the University of Birmingham's home source 'Rigaku HighFlux HomeLab X-ray diffraction system'. ESRF refers to the European Synchrotron Radiation Facility, Grenoble, France.

	Processing							
Dataset Identifier	*Resolution Range (Å)*	*No. Observations*	*No. Unique*	*Multiplicity (%)*	*Completeness (%)*	*I/σI*	R_{sym} *(%)*	*Mosaicity*
NfsA Succinate 1	46.47-1.16 (1.23-1.16)	975,954 (103,337)	67,619 (10,081)	14.4 (10.3)	98.4 (94.8)	35.7 (6.6)	4.3 (35)	0.169
NfsA Succinate 2	46.48-1.24 (1.31-1.24)	416,204 (46,685)	59,466 (8,483)	7.0 (5.5)	95.4 (84.9)	30.8 (5.3)	3.8 (23)	0.157
NfsA Fumarate	46.59-1.25 (1.31-1.25)	512,709 (30,520)	57,570 (6,571)	8.7 (4.6)	93.5 (73.5)	26.1 (3.7)	4.3 (40)	0.193
NfsA Empty	46.54-1.10 (1.16-1.10)	243,913 (21,349)	80,806 (10,160)	3.0 (2.1)	90.4 (70.9)	8.21 (1.5)	2.5 (6.2)	0.133
NfsA Ethylene Glycol	46.26-1.50 (1.58-1.50)	342,588 (22,629)	33,331 (4,769)	10.3 (4.8)	95.2 (85.7)	40.8 (5.5)	3.6 (27)	0.219
NfsA FMN	46.71-1.03 (1.09-1.03)	844,480 (25,188)	105,128 (7,409)	8.0 (3.4)	96.2 (91.8)	27.9 (2.1)	4.4 (58)	0.148
NfsA Quinol	46.46-1.25 (1.32-1.25)	379,094 (25,991)	59,285 (7,494)	6.4 (3.5)	97.6 (84.9)	19.2 (2.4)	5.2 (55)	0.256
NfsA Quinone	46.59-1.70 (1.79-1.70)	273,625 (16,275)	23,824 (3,102)	11.4 (5.3)	98.2 (88.7)	26.8 (4.6)	9.1 (35)	0.123
NfsA Nitrofurantoin	46.38-1.09 (1.15-1.09)	283,824 (21,970)	82,863 (9,425)	3.4 (2.33)	91.4 (64.5)	19.4 (3.7)	3.7 (22)	0.142

Table 5-3: Processing statistics of NfsA datasets. Values in parentheses are for the highest resolution shell. $R_{sym} = (\Sigma_{hkl}\Sigma_j|I_{hkl,j} - \langle I_{hkl}\rangle|)/\Sigma_{hkl}\Sigma_j I_{hkl,j}$.

	Refinement						Mean B Factors (Å^2)			
Dataset Identifier	*No. of Reflections*	*No. Water Molecules*	*R factor (%)*	*R_{free} (%)*	*RMSD bond angles (°)*	*RMSD Bond Lengths (Å)*	*Protein*	*Ligand*	*FMN*	*Water*
NfsA Succinate 1	142,450	327	11.06	13.01	1.569	0.015	15.9	10.8	6.7	27.1
NfsA Succinate 2	59,481	389	14.96	16.27	1.054	0.005	14.7	11.2	6.6	28.7
NfsA Fumarate	57,566	381	14.57	15.90	1.062	0.005	12.8	9.0	6.8	26.7
NfsA Empty	80,944	416	12.69	14.13	1.469	0.010	13.3	n/a	5.9	25.8
NfsA Ethylene Glycol	33,483	305	14.31	17.02	1.178	0.008	12.8	18.3*	6.4	27.0
NfsA FMN	101,426	404	11.43	13.88	1.332	0.009	13.9	22.6	6.1	25.7
NfsA Quinol	59,267	410	15.14	17.3	1.043	0.005	13.9	27.0	6.3	29.5
NfsA Quinone	23,814	293	14.95	18.31	0.976	0.006	13.3	25.3	5.5	25.6
NfsA Nitrofurantoin	83,276	394	11.67	13.76	1.709	0.009	17.7	14.4	6.7	27.4

Table 5-4: Refinement statistics for NfsA datasets. * This is an average of two conformations of two ethylene glycol molecules modelled in the active site. *R-factor* = $(\Sigma| |F_{obs}|-|F_{calc}||)/\Sigma|F_{obs}|$.The *R-free* is the R-factors calculated from the 5% of data left out of refinement.

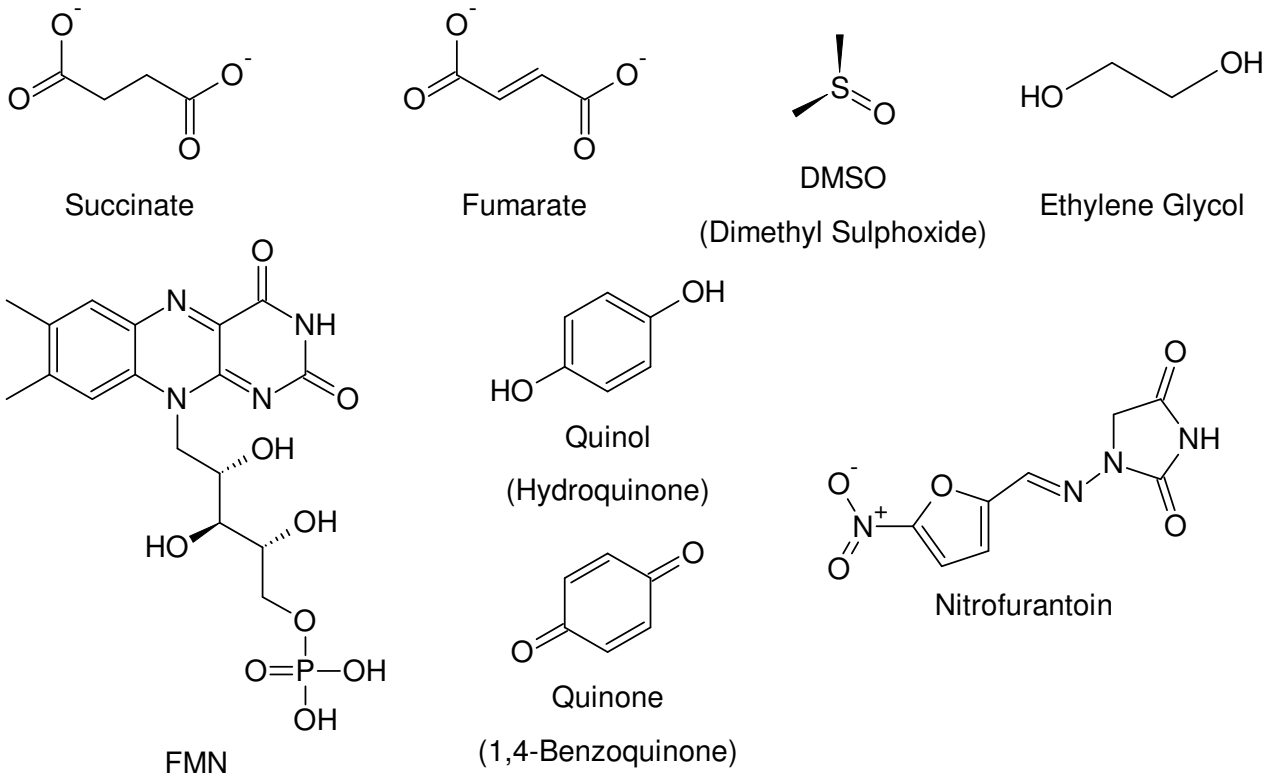


Figure 5-3: The chemical structures of the ligands found in the NfsA structures presented in this chapter. Figure produced using ChemSketch (ACD/labs).

All the structures show low R-factors and B-factors, and the majority of the density was very clear. All the ligands have full occupancy, with the exception of the quinol, which was 50:50 in two conformations, and the second FMN where the atoms were each individually refined for occupancy in PHENIX (Adams et al., 2002) and ranged from 50% to 100%. Ligand B-factors are lower than the protein average for all the FMN prosthetic groups and the succinate, fumarate and nitrofurantoin ligands. The B-factors for the other ligands are less than twice as high as the protein average. All the structures have good statistics and the active sites are clear.

5.5 Overview of NfsA Structures

In common with NfsB, NfsA exists as a homodimer, with two FMN containing active sites located at the dimer interface. The structure is shown in Figure 1-28, where the two halves of the dimer can be seen, and described in Section 1.6.3.2. The only previously published crystal structure of NfsA (Kobori *et al.*, 2001) was processed in a triclinic space group and contained one copy of the dimer in the ASU. The structure presented here was of a form previously observed in this group (Jarrom, 2008), a monoclinic space group containing a single monomer in the ASU. The dimer is formed through crystallographic interactions with an adjacent unit cell. A comparison of the structures presented in this chapter (except NfsA FMN which will be discussed later) with the previous structure is shown in Figure 5-4. The only significant variation in backbone position is seen in the loop between helices G and H (residues 197-211), a loop which is always poorly modelled in the structures in this chapter and has high B-factors (generally at least twice the average for the protein). The published structure (green in Figure 5-4) has a different loop conformation from that of all the other structures determined here. There are no differences in the side chain conformations of the active site residues between any of the structures.

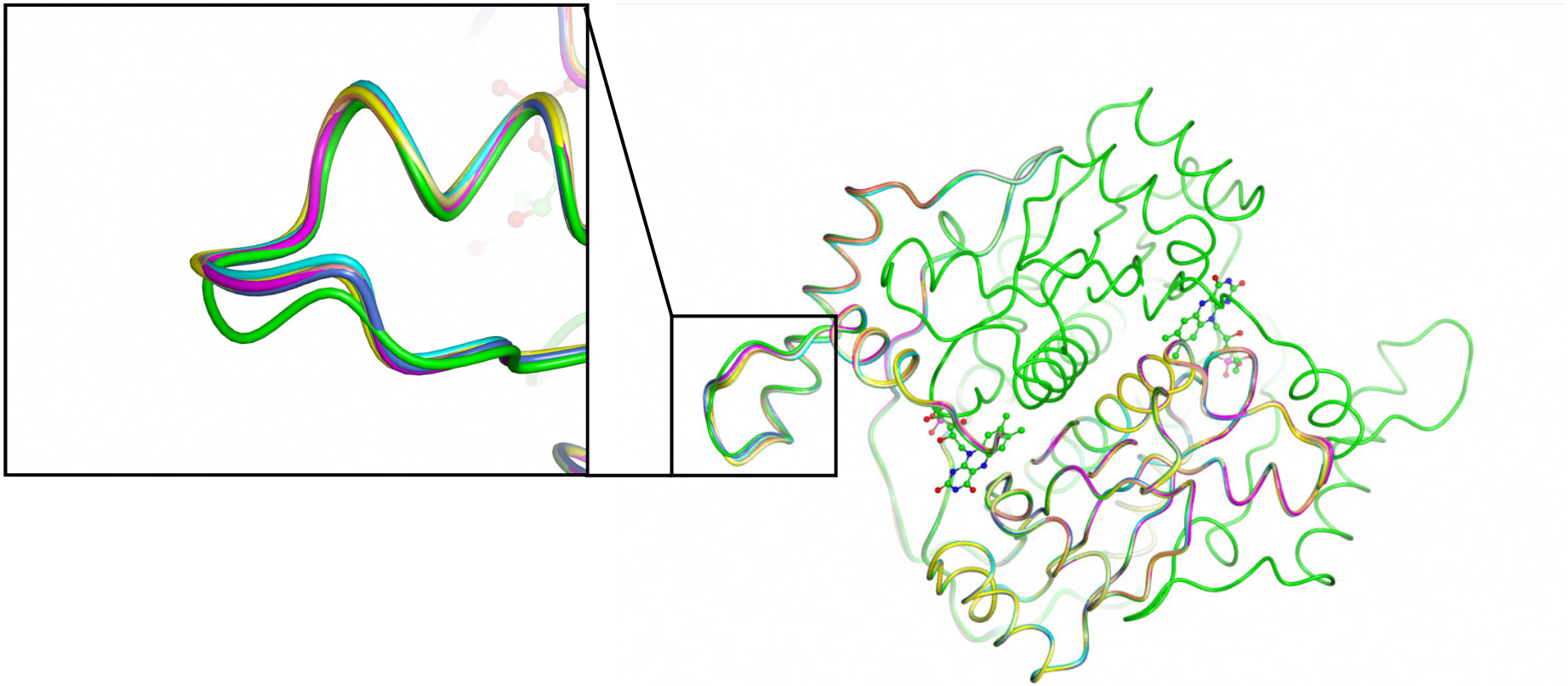


Figure 5-4: An overlay of the smoothed Cα traces of the published structure of NfsA (green - pdb I.D. 1F5V (Kobori *et al.*, 2001)), and all the structures determined in this chapter (except the NfsA FMN structure). The published structure is a dimer, whilst all the structures from this chapter are monomers and so have been aligned to one half of the dimer using THESEUS (Theobald and Wuttke, 2008). The second half of the dimer of the published structure aligns perfectly with the first. The insert shows the flexible loop between helices G and H. Figure produced using CCP4 Molecular Graphics.

5.6 NfsA Crystallized in the Presence of Succinate and Fumarate

As has been mentioned earlier, our group had previously solved the structure of NfsA showing density in the active site that resembled succinate or fumarate (both found to be inhibitors) (Jarrom, 2008). Because of this the decision was made to add first succinate and then fumarate to the crystallization solution. In both cases a succinate-like molecule was found bound in the active site of NfsA above the flavin ring system as shown in Figure 5-5. One of the carboxylates of the succinate forms a salt bridge with the guanidinium group of Arg225, with the two NH_2 groups lining up with the oxygen of the succinate. The guanidinium of Arg 133 is also in a position to hydrogen bond this carboxylate at a similar distance. The other carboxylate hydrogen bonds to both the backbone NH of Ser 41 and to the O2' alcohol of the FMN ribityl. The central bond, which is oxidised in the conversion of succinate to fumarate (and reduced in the reverse reaction) is seem directly above the reactive N5 nitrogen of the FMN (the *si* face - Figure 1-4).

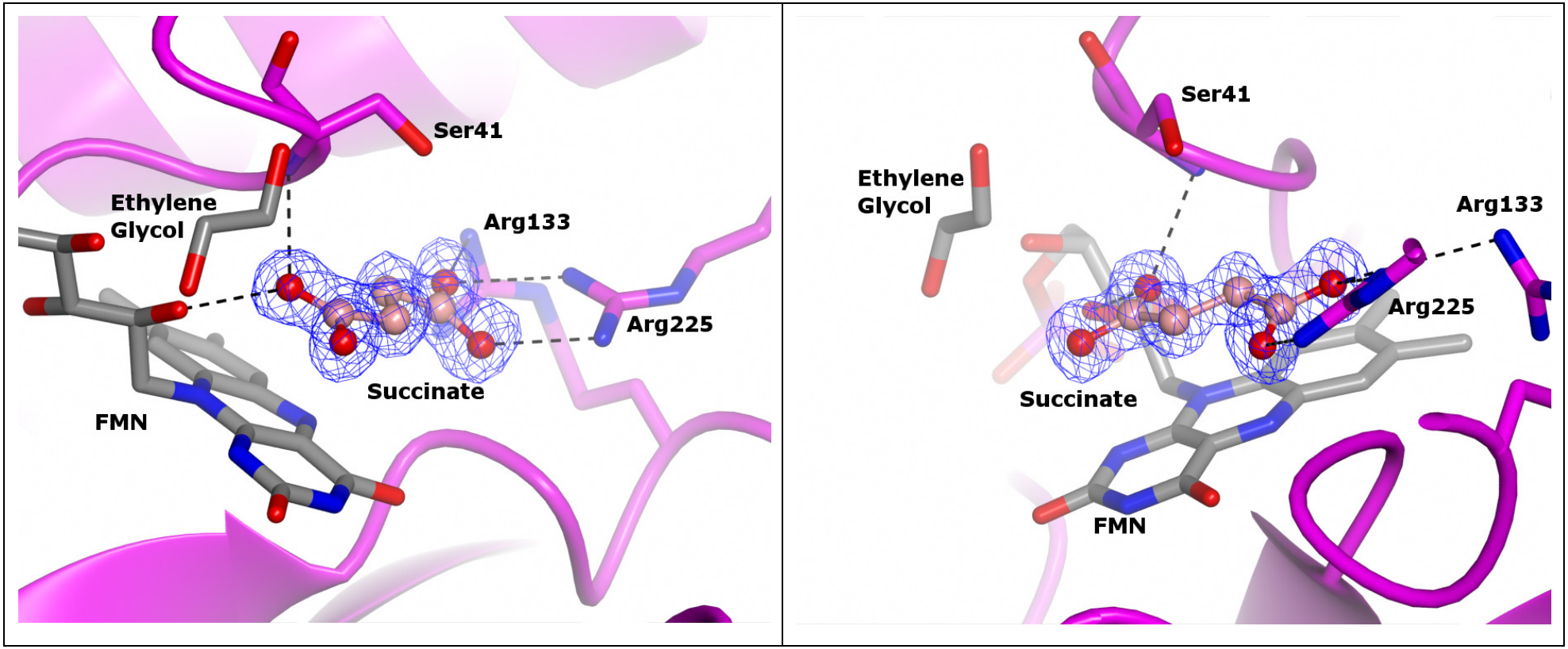


Figure 5-5: Succinate bound in the active site of the NfsA Succinate 1 structure. The two images are separated by 90 °. The succinate is shown in pink. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.5 σ) has been clipped to within 2.5 Å of the ligand. Figures were produced using CCP4 Molecular Graphics.

However, initial crystals grown in the absence of succinate (or fumarate), also showed identical density in the active site. In fact NfsA co-crystallized with malonate showed density indistinguishable from NfsA co-crystallized with either succinate or fumarate (Figure 5-6 c)-e)). Similar density was also seen when NfsA was co-crystallized with L-ascorbic acid, histidine and tryptophan (data not shown), which were added as part of an unpublished additive screen aiming to improve crystallization. All the structures had low R-factors and the B-factor for the ligand was always lower than the average for the protein. It was only when the precipitant was changed from PEG 3,350 (Sigma) to PEG 3,000 (Fluka) that the succinate-like density was no longer seen in the active site (the active site contained ethylene glycol, see section 5.7). The density was only then regained by co-crystallizing with succinate or fumarate. This implied that the original succinate-like density was as a result of the contaminant from the PEG 3,350 purchased from Sigma.

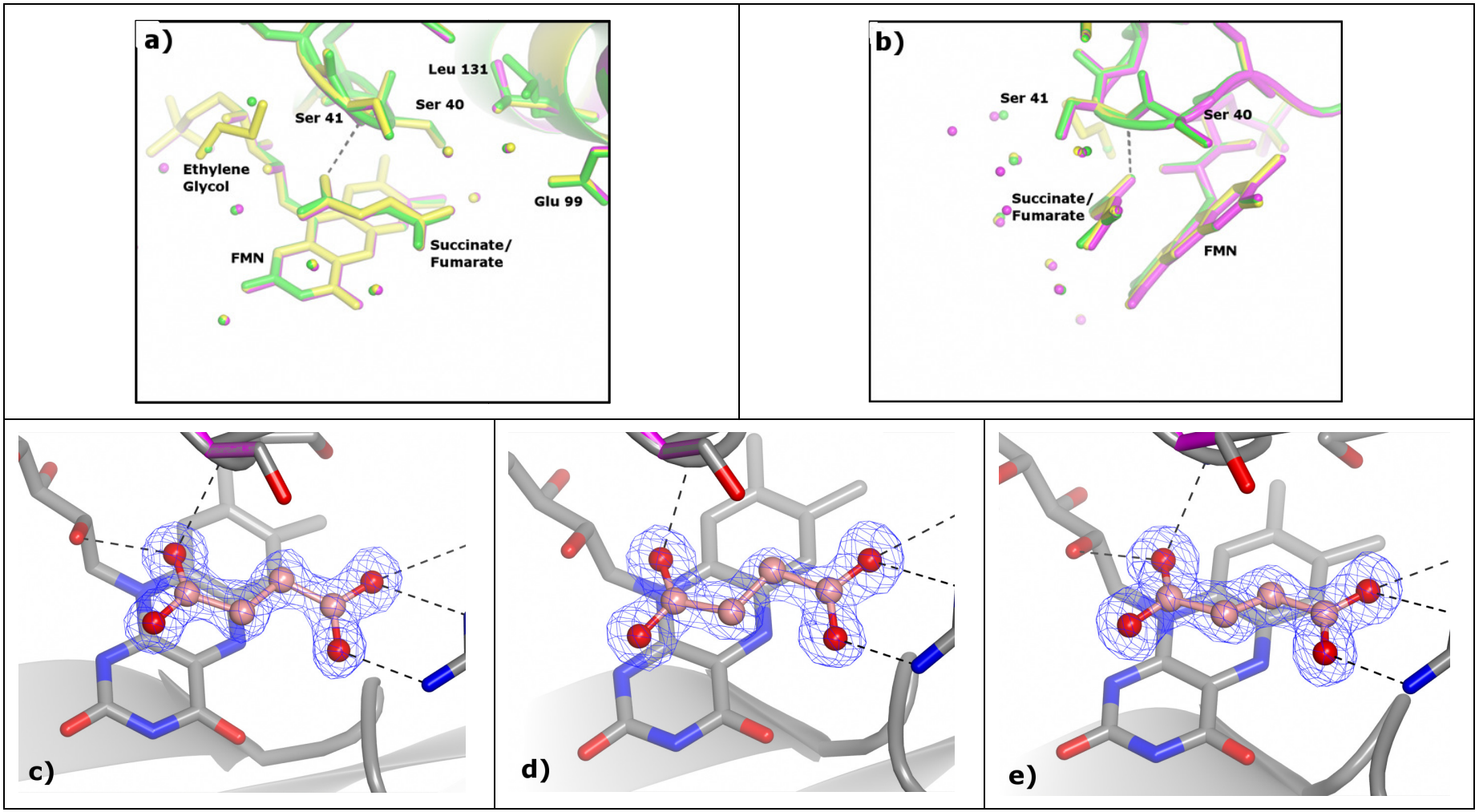


Figure 5-6: a) and b) A comparison of the active sites of the NfsA Succinate 1 (yellow), Succinate 2 (purple) and Fumarate (green) structures aligned using THESEUS (Theobald and Wuttke, 2008). Two images are 90 ° apart. Electron density of the active site ligand in c) NfsA Succinate 1, d) NfsA Succinate 2 and d) NfsA Fumarate. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

5.7 NfsA bound to ethylene glycol

When the PEG was changed to PEG 3,000 (Fluka), short soaks (1-2 min steps of 5 %, 10 % then 15 %) with ethylene glycol were sufficient to cryo-protect the protein crystals and gave a structure with nothing above the FMN in the active site. With longer soaks (5-10 min steps) ethylene glycol was seen in the active site. Two alternative conformations of two ethylene glycol molecules were found above the flavin ring system, as shown in Figure 5-7. One ethylene glycol molecule mirrors the interaction of the succinate/fumarate carboxylate with the guanidinium group of Arg 225, only forming hydrogen bonds rather than ionic interactions. The second ethylene glycol interacts with either the backbone nitrogen of Ser 41 or the O2' alcohol of the FMN ribityl group, mirroring the interactions of the other carboxylate of succinate/fumarate (Section 5.6). It is possible that the solvent plays some part in the alternative conformations seen in this structure.

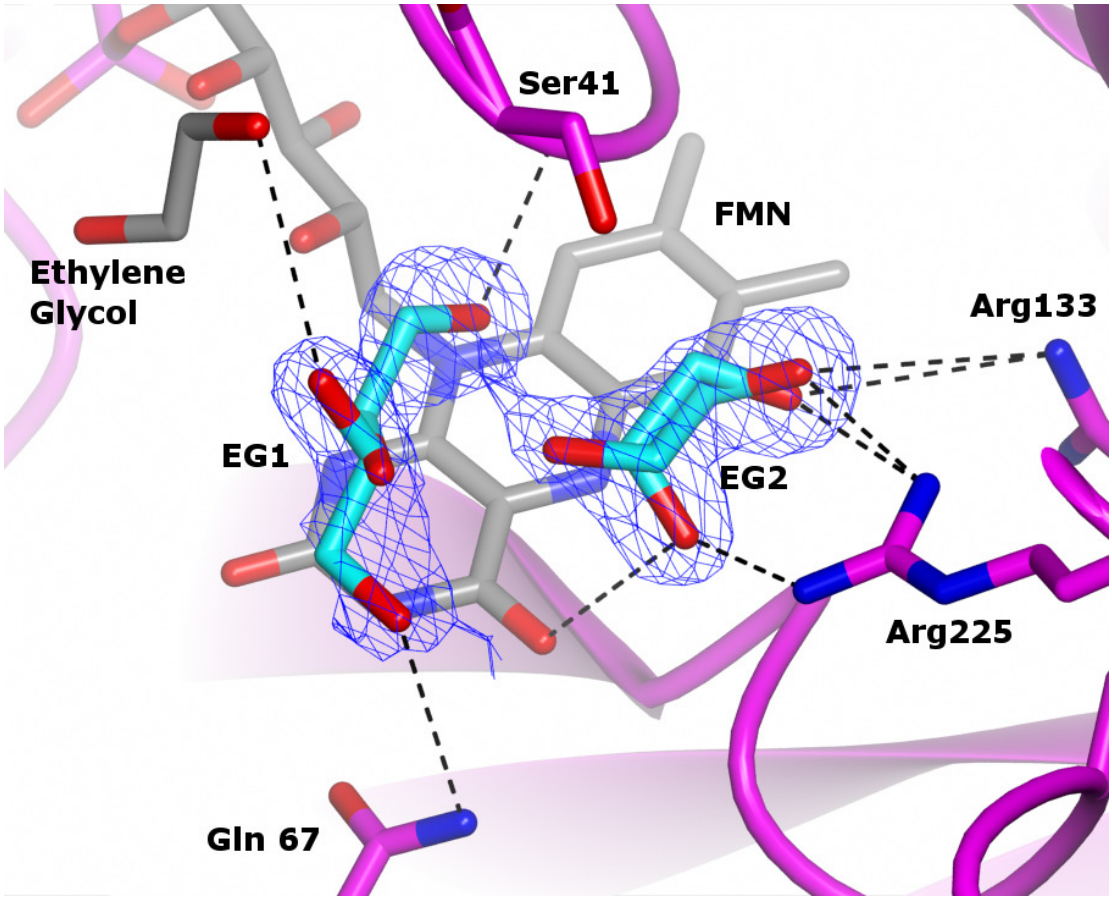


Figure 5-7: Active site of the NfsA Ethylene Glycol. The two alternate conformations of the two active site ethylene glycols (labelled EG1 and EG2) are shown in light blue. The FMN and another ethylene glycol are shown in grey. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the ligands. Figure was produced using CCP4 Molecular Graphics.

In order to obtain complexes with any substrates ethylene glycol could not be used as a cryo-protectant as it bound in the active site. To prevent this DMSO, was used as a cryo-protectant, in which role it could double as a solvent for substrates including nitrofurantoin. Using DMSO as a cryo-protectant gave a structure with only water in the active site (data not shown). Glycerol and several sugars were also tested as cryo-protectants but crystals dissolved or did not diffract.

5.8 NfsA bound to a second FMN

An NfsA crystal soaked with hydroquinone showed electron density in the active site that did not fit hydroquinone, but showed at least 2 fused aromatic rings. The only molecule in the crystallization mixture with more than one ring was the FMN, and the density seemed to suggest that ribityl and phosphate were present, so a second copy of the cofactor was modelled into the active site. The flavin fit the density well as shown in Figure 5-8. The ring system's density is all in a plane and there is clear density for each atom of the ribityl tail. The density for the isoalloxazine ring system is clear enough that the holes in the centre of each ring can be clearly seen. The R-factors of the structure are low and the average B-factor for the FMN is around twice the average for the protein, although still not high.

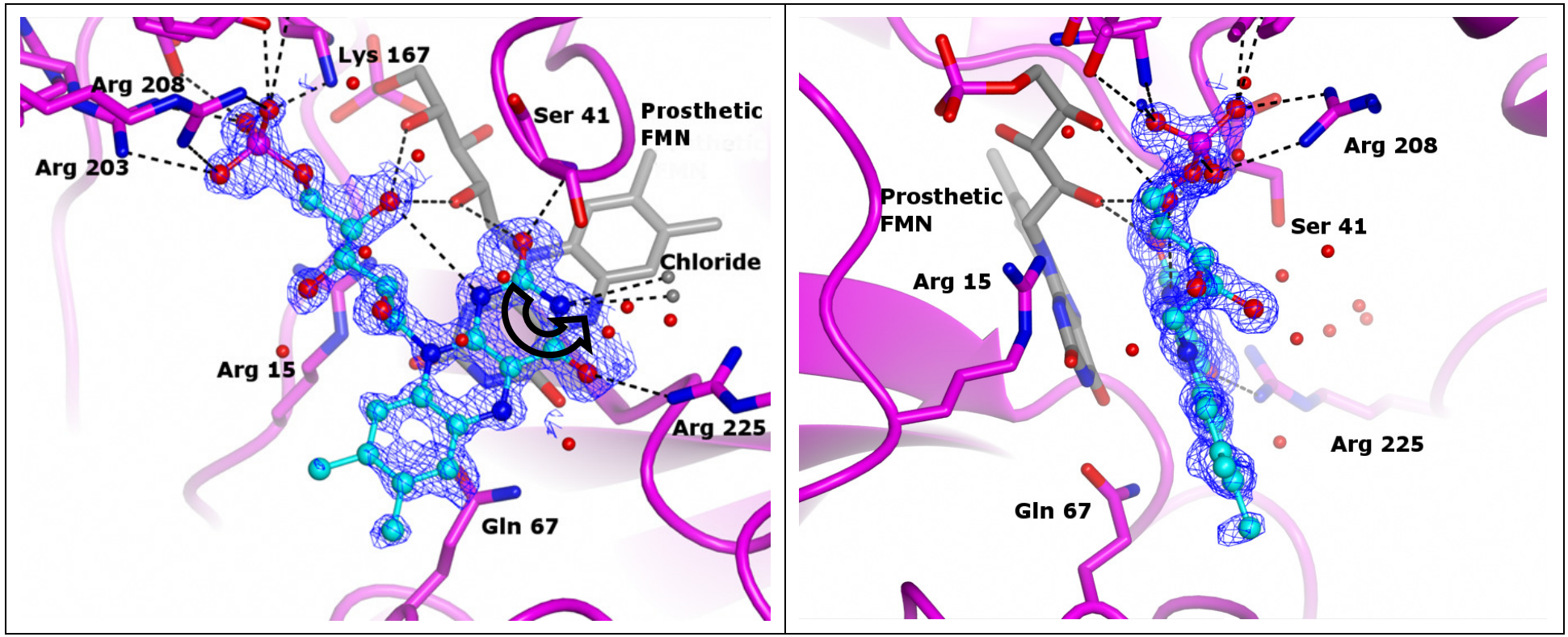


Figure 5-8: Active site of NfsA FMN. Two images are separated by 90 °. The second 'substrate' FMN is shown as a light blue ball and stick model. The FMN prosthetic group is shown as grey cylinders. The black arrow on the left hand image shows the suggested direction of rotation around the N3 nitrogen. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

FMN has previously been shown to inhibit NfsA (but not NfsB) (Zenno *et al.*, 1996a). The second, inhibitor FMN is bound so that its isoalloxazine ring system is stacked above the prosthetic FMN's ring system, but is displaced so that only two rings overlap. In comparison to the prosthetic FMN, the second FMN is rotated 180 ° around an axis perpendicular to the N10-C1* bond (separating the ring system and the ribityl tail). This leaves the second flavins pyrimidinedione ring above the centre ring of the prosthetic group, and the central ring above the dimethylbenzene ring. The electron density is stronger towards the centre of the active site, i.e. at the pyrimidinedione ring, and weakest towards the outside of the enzyme at the dimethylbenezene ring. There are areas of positive (unmodelled) density all around the FMN ligands, and the second FMN was refined for individual atom occupancy showing that atoms further from the centre of the active site had lower occupancy, suggests that the inhibitor-like FMN molecule exists in several orientations effectively performing a rigid anti-clockwise rotation around an axis perpendicular to the ring system roughly through the N3 nitrogen (show with a black arrow in Figure 5-8). The electron density is all in a plane showing that the second FMN does not display the butterflying of the isoalloxazine ring system that is seen with the FMN prosthetic group.

The second FMN forms a series of H-bonds to the O2' and O4' alcohol groups of the prosthetic FMN ribityl, through its own O4' alcohol and the O2 carbonyl oxygen. The two carbonyl oxygens of the isoalloxazine ring hydrogen bond to the protein; the O2 to the backbone nitrogen of Ser 41 and the O4 to the guanidinium of Arg 225. The binding of the second FMN's phosphate causes the only variation in secondary

structure seen with any of the structures shown in this chapter. The poorly modelled loop between 197 and 211 has moved to create a phosphate-binding pocket. This shift is compared to all the other structures in Figure 5-9. In the NfsA FMN-bound structure the loop has two alternate conformations. The occupancies were 50:50 but the B-factors would suggest that there is a dominant conformation (conformation A), which is the new one, and a secondary conformation (conformation B), that is identical to that seen in all the other structures. The second FMN's phosphate atoms have ~50% occupancy, further suggesting that the driving force behind the loop movement is the formation of the phosphate binding pocket.

To remove model bias the loop was removed from refinement and an omit map (i.e. density calculated from refinement from a model lacking the loop residues) produced before both new and old loop conformations were remodelled. The distances between these two conformations and the average B-factors of these residues are detailed in Table 5-5. The main chain B-factors (Table 5-5) are lower for the new conformation (A – mostly between 9 and 11.5 $Å^2$) than the old (B – ranging from 8.8 to 47.3 $Å^2$). Most other structures have average B-factors for the loop between 15 and 30 $Å^2$, higher than the new loop in the FMN-bound structure, but lower than the old.

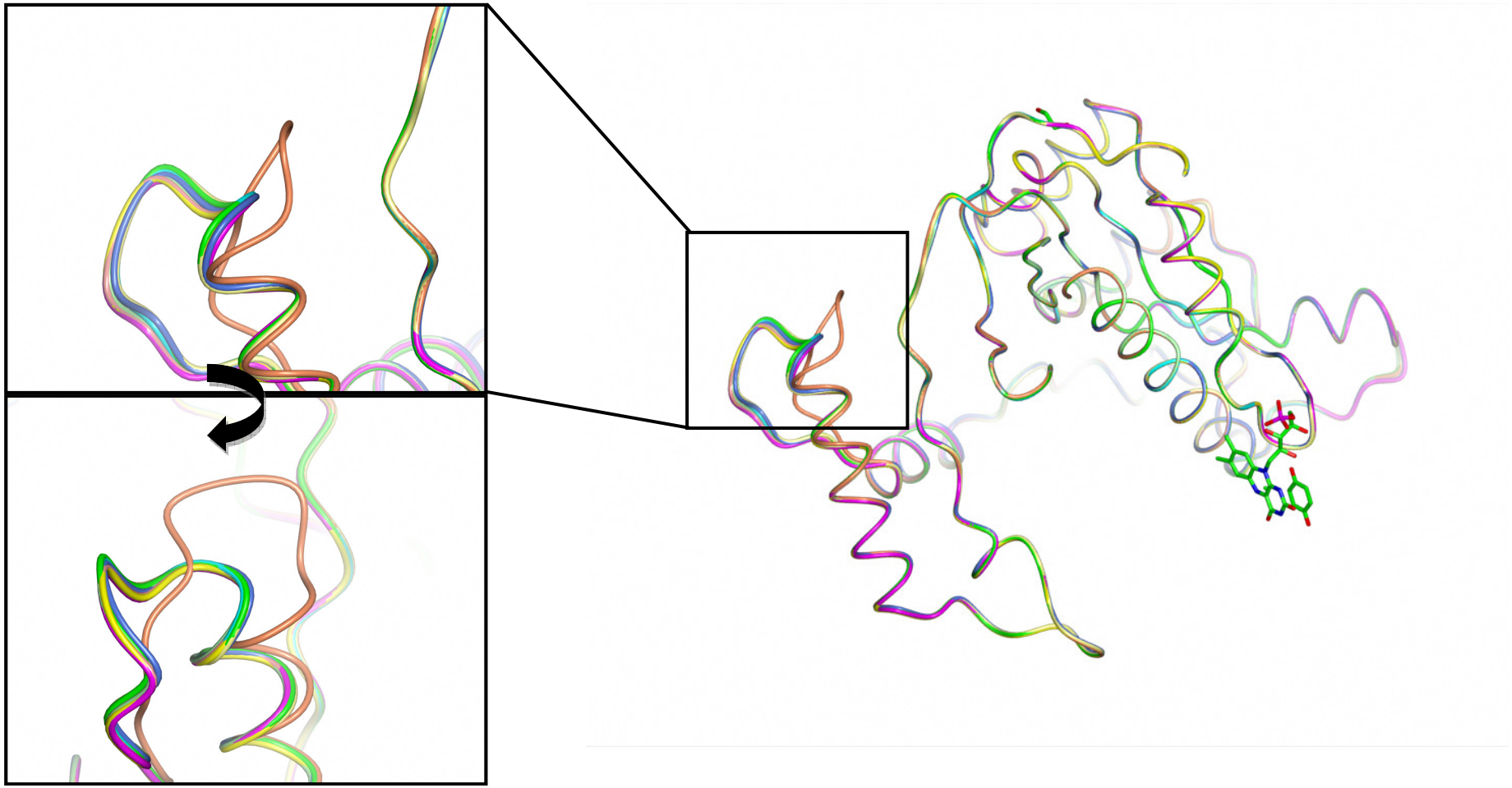


Figure 5-9: An overlay of the smoothed Cα traces of all the monomer structures detailed in this chapter. Structures were aligned using THESEUS (Theobald and Wuttke, 2008). The dominant conformation of the flexible loop in the NfsA FMN structure is shown in orange. The insert shows the flexible loop, the two images are separated by 90 °. Figures were produced using CCP4 Molecular Graphics.

Residue	Cα(A) - Cα(B) Distance (Å)	Average B-factor – ($Å^2$) Conformation A	Average B-factor – ($Å^2$) Conformation B
Ala 197	0.5	10.1	12.3
Glu 198	1.0	10.0	15.0
Tyr 199	1.5	9.5	13.1
Tyr 200	1.7	9.0	13.8
Leu 201	2.2	11.4	24.3
Thr 202	3.7	13.2	25.5
Arg 203	6.2	10.2	17.8
Gly 204	10.8	11.5	16.5
Ser 205	10.7	10.4	28.1
Asn 206	11.0	11.3	44.9
Asn 207	9.5	11.0	47.3
Arg 208	7.2	10.5	31.4
Arg 209	3.8	12.6	15.3
Asp 210	1.2	11.0	14.0
Thr 211	0.5	9.3	8.8

Table 5-5: Comparison of the two conformations of the NfsA flexible loop in the NfsA FMN structure. The dominant conformation (A) is the changed conformation seen only in this structure. The distance between the α-carbons of each conformation is given as is the average B-factor for main chain atoms of the residue. The average for the whole protein is 9.1. The resolution of the structure is 1.09 Å, and the maximum likelihood based coordinate error is 0.1 Å.

The maximum shift between equivalent Cα's is 11 Å and this movement seems to be driven by the creation of a phosphate-binding site for the second FMN. The binding site and change in position of the residues involved is shown in Figure 5-10. The Arg 203 and Arg 208 guanidinium groups have moved to hydrogen bond to the phosphate oxygens, as has Tyr 200, but Tyr 199 moves so that it does not. There is a change in

salt bridging with the salt bridge between Lys 167 and the FMN cofactor being replaced by a salt bridge between Lys 167 and the substrate FMN, requiring a 2.9 Å shift in the position of the Lys 167 NZ. The side chains of the residues which form the phosphate binding pocket have lower B-factors than those that do not.

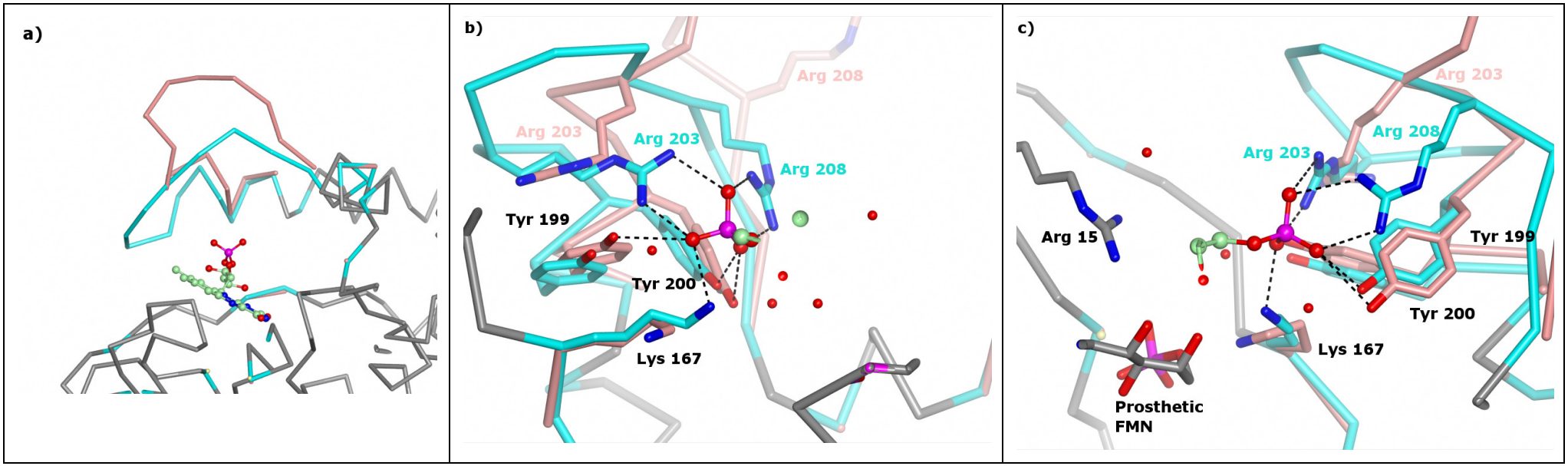


Figure 5-10: The flexible loop change in the NfsA FMN structure. The new loop conformation (A) is shown in light blue and the old (B) in pink. The second FMN is shown with carbon atoms in green. Image a) shows the overall loop change, while images b) and c) show the phosphate binding pocket created and are 90 ° apart. Figures produced using CCP4 Molecular Graphics.

While quinol was soaked into this crystal it was not seen in the active site, however there was density that appeared to match quinol near two surface cysteines, Cys 9 and Cys 90, neither of which are near the active site. Initial attempts to model quinol into the density led to the model being forced away from the density during refinement. If the model was fitted to the density with no restraints the nearest carbon atom of the quinol to the cysteine sulphur was actually closer than the β-carbon bonded to the sulphur, suggesting a covalent adduct. Quinone can modify surface cysteines of proteins forming S-cysteinyl-benzoquinol adducts (Mason and Liebler, 2000) (Figure 5-11), and, while it was quinol in the soak, it is possible that some of the quinol could be converted to quinone in the presence of the crystallization solution or free radicals produced by X-ray damage. Of the two surface cysteines in NfsA that showed extra density in was clearest with Cys 90 (Figure 5-12), which was refined with full occupancy and had an average B-factor of 11.5 $Å^2$ (protein average = 13.9 $Å^2$). This density was only seen with crystals grown in presence of quinone or quinol.

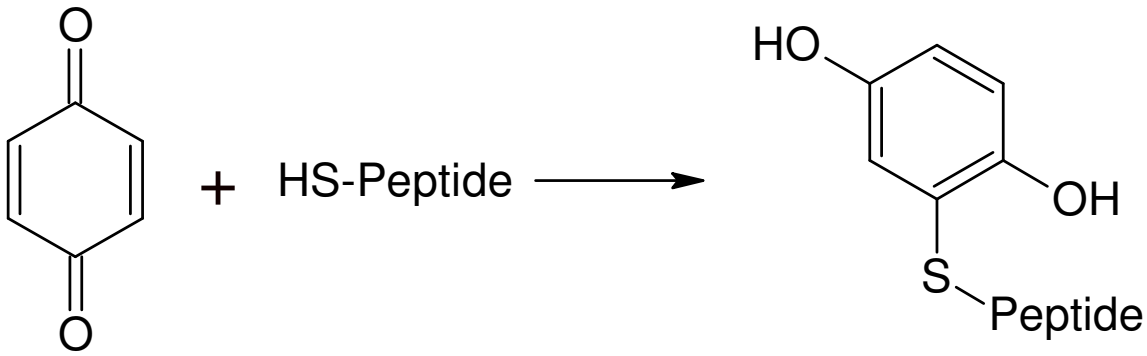


Figure 5-11: The reaction of quinone with surface peptide cysteinyl thiols forming S-cysteinyl-benzoquinol adducts. Figure adapted from Figure 1 in reference (Mason and Liebler, 2000). Figure produced using ChemSketch (ACD/labs).

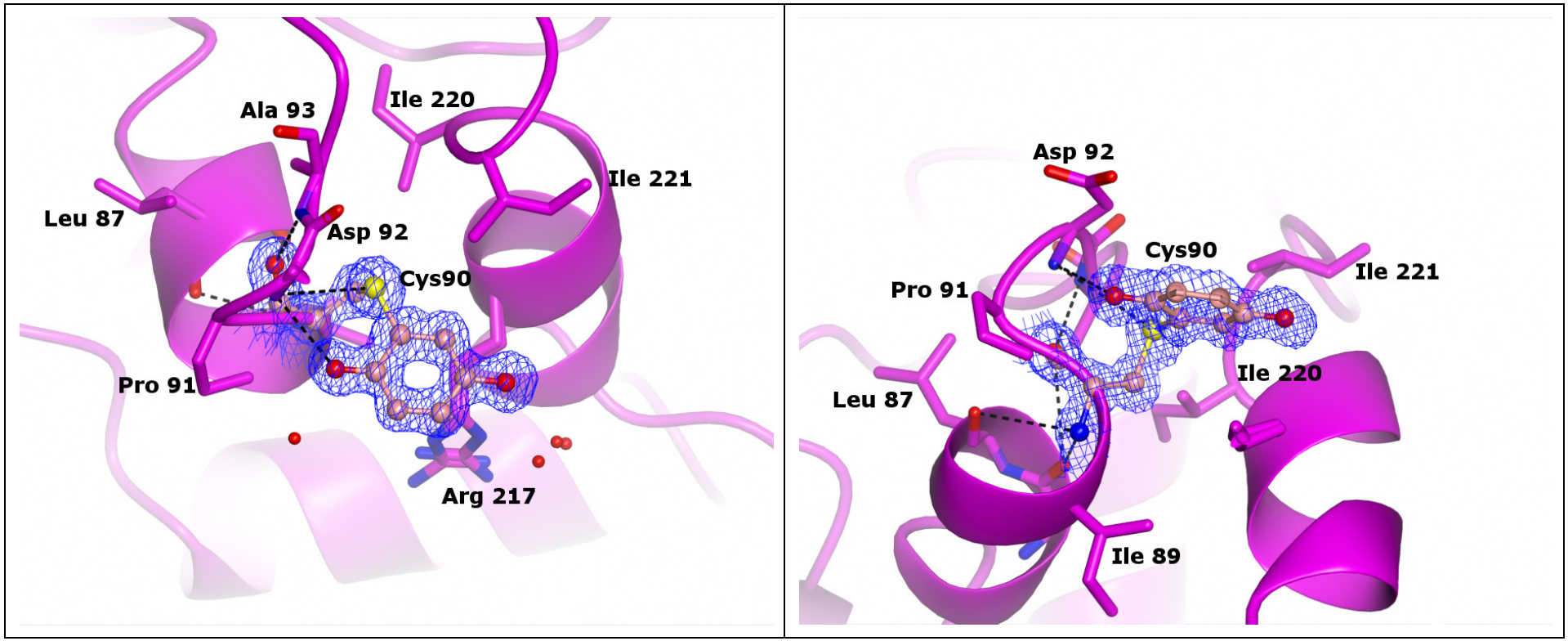


Figure 5-12: Modification of Cys 90 in the NfsA FMN structure showing the new S-cysteinyl-benzoquinol adduct (pale pink). Two images are 90 ° apart. H-bonds are represented by black dashed lines. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.0 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

The second residue (Cys 9) showed further density in addition to the quinol adduct, indicating an addition of a 5-membered ring to the quinol. The only chemical in the crystallization solution containing such a ring is the imidazole used as a buffer. Imidazole can be added to benzoquinones (Escolástico *et al.*, 1994) (Figure 5-13), and in such a way that a single addition would match the density seen next to Cys 9. An adduct of both quinol and imidazole has been modelled to Cys 9 in the FMN-bound NfsA structure (Figure 5-14). The residue was refined for occupancy of individual atoms, showing the cysteine at 100 %, the cysteinyl-benzoquinol at 70 % and the whole adduct at 45 %, giving an indication of the levels of each stage of modification.

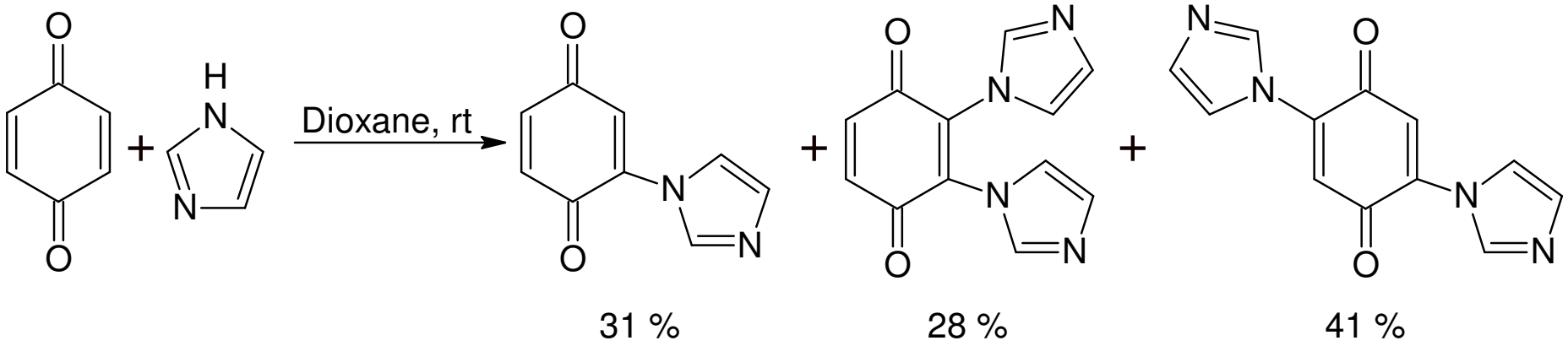


Figure 5-13: Reaction of quinone with imidazole in dioxane at room temperature. Proportions of products given are those detailed by Escolástico *et al.*, figure adapted from reference (Escolástico *et al.*, 1994). Figure produced using ChemSketch (ACD/labs).

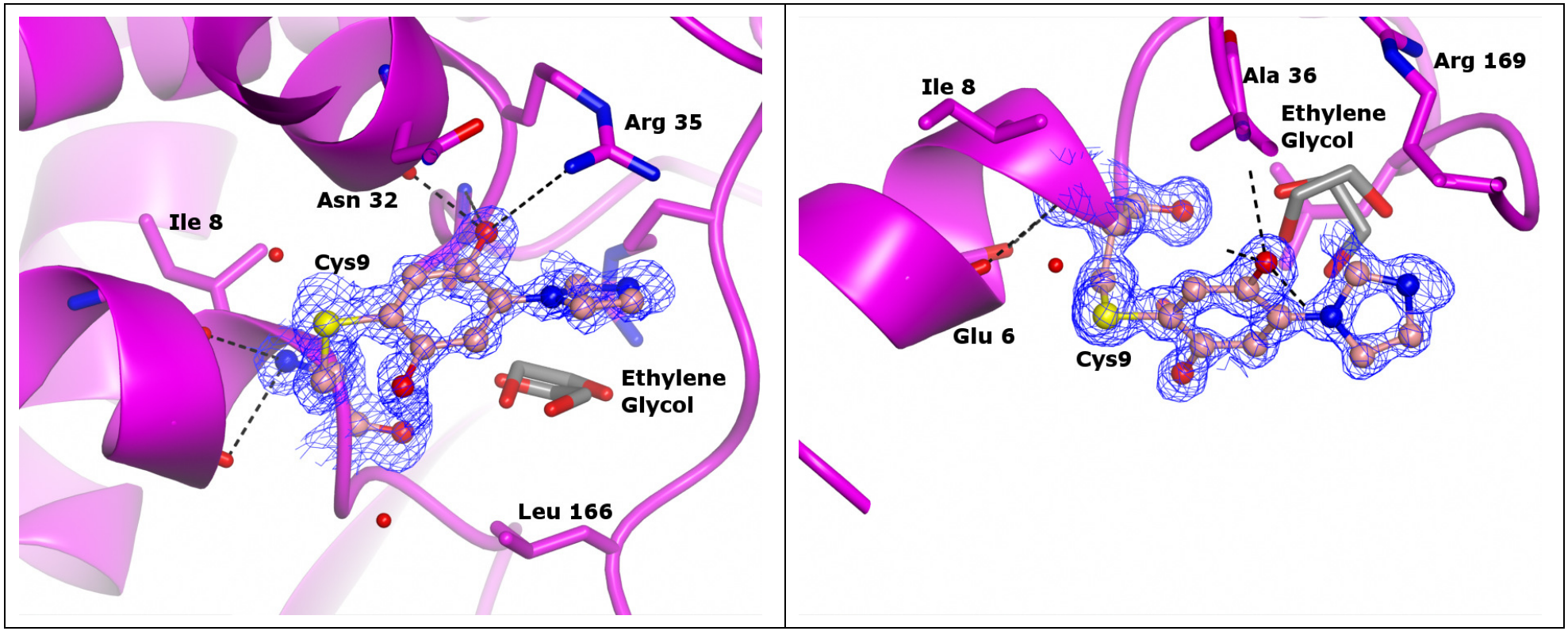


Figure 5-14: Modification of Cys 9 in the NfsA FMN structure showing the new adduct (pale pink). Two images are 90 ° apart. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~2.0 σ) has been clipped to within 2.5 Å of the residue and adduct. Figures were produced using CCP4 Molecular Graphics.

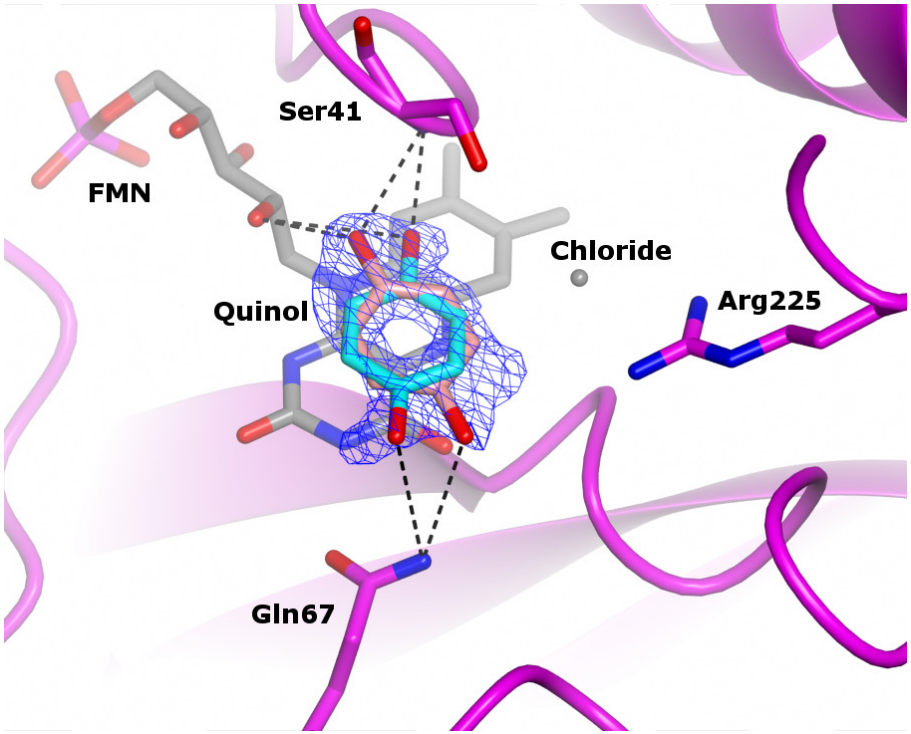
FMN was added to the protein solution to 20 μM after purification to ensure full occupancy of the prosthetic group and thus fully active enzyme. After the PEG was changed, FMN was seen in the active site. So for further experiments excess FMN was dialysed out of the protein solution prior to crystallization.

5.9 NfsA bound to quinol and quinone

NfsA catalyzes the reduction of quinone, to give quinol (see Figure 5-3 for structures). NfsA was co-crystallized with both substrate and product, with both returning similar density in the active site (Figure 5-15). The density suggested that quinone could rotate in the active site. As mentioned in Section 5.8 the is is possible that some or all of the quinone could be converted to quinol in the X-ray beam, and so both structures may be quinol. This rotation is around a point in the centre of the benzene ring, within a plane roughly parallel to the flavin ring system. One carbonyl oxygen (or alcohol) of the quinone (quinol) interacts with a combination of the FMN O2' alcohol and the backbone nitrogen of Ser 41, whilst the other interacts with the guanidinium of Arg 225 and/or the amine of Gln 67. In the quinol-bound NfsA structure the ligand has been modelled in two conformations about this rotation point. The same modification of Cys 90 as the NfsA FMN structure (but not Cys 9) was seen with the quinol soak, but not with the quinone soak (although some extra density was seen), even though the quinone is the reactive species. This could be due to the much shorter soak times used with quinone, and lower concentration of the ligand. Both ligands have full occupancy

(50:50 for the two quinol conformations) and B-factors twice the average for the protein.

a)



b)

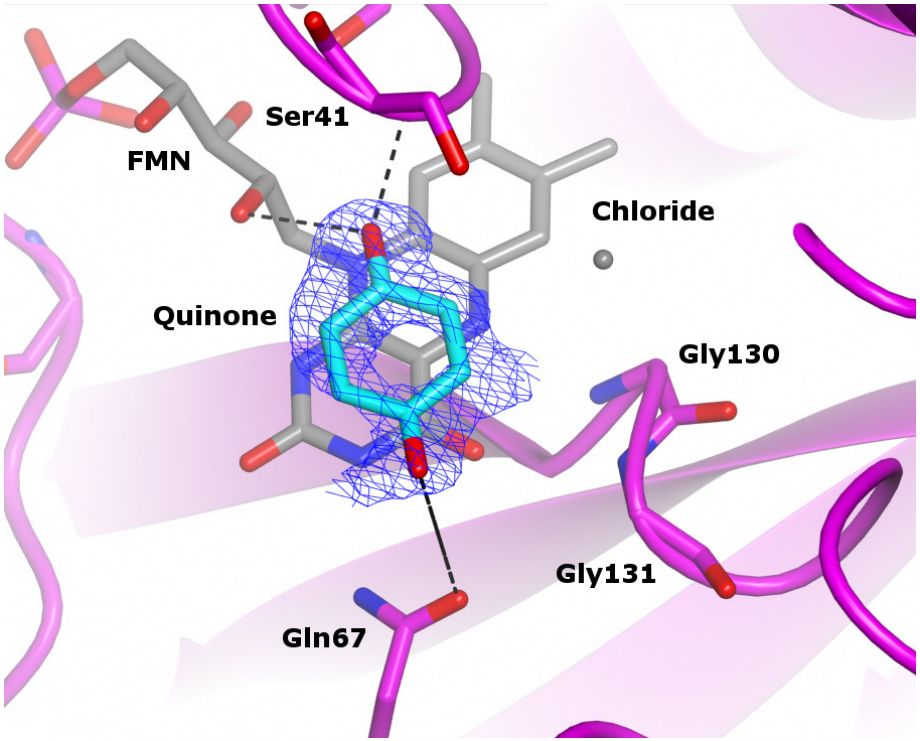


Figure 5-15: The active sites of a) NfsA Quinol and b) NfsA Quinone. Two conformations of quinol are shown in light blue and light pink in a) and one of quinone in light blue in b). The FMN is shown as grey. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

5.10 NfsA bound to nitrofurantoin

The only successful attempt to obtain an NfsA complex with a nitroaromatic substrate was with the antibiotic nitrofurantoin. Figure 5-16 shows the electron density in the active site and the binding of nitrofurantoin. The density was clear, and the ligand had full occupancy and its B-factor was lower than the average for the protein. The ligand was bound with the nitro group bound away from the FMN ring system and the reactive centre of the N5 nitrogen (see Figure 1-4 c) for numbering). The imidazole ring rather than the nitrofuran ring is above the centre of the flavin ring system, with the ketone oxygens of the ligand hydrogen bonding to the backbone nitrogen of Ser 41 and the guanidinium of Arg 225. The nitrogen of the imidazole ring is in the active position above the FMN N5. In addition to the aromatic stacking interactions of the rings, the alcohol groups of the FMN ribityl moiety also hydrogen bond to various points on the ligand (see Figure 5-16).

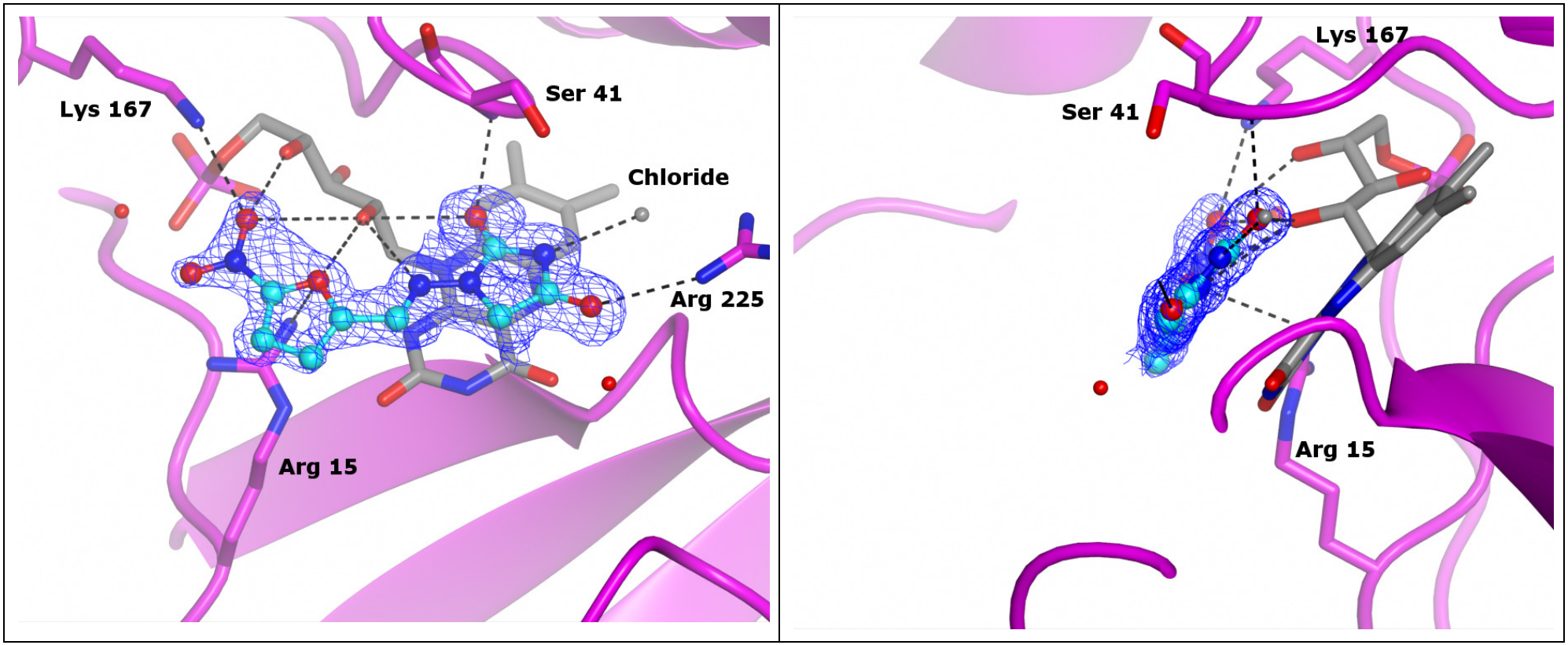


Figure 5-16: Active site of NfsA Nitrofurantoin. Two images are separated by 90 °. The FMN is shown in grey and the nitrofurantoin in light blue. H-bonds are shown as black dashed lines. The σ_A-weighted $2F_o$-F_c electron density map (contoured to ~1.5 σ) has been clipped to within 2.5 Å of the residue. Figures were produced using CCP4 Molecular Graphics.

5.11 Steady-State Kinetics

5.11.1 Succinate Inhibition

As succinate was found in the active site of the crystal structure of NfsA, the effect of sodium succinate on the NfsA-catalysed reduction of NFZ with NADPH has been investigated. The steady-state kinetics of the reduction of NFZ with NADPH have been described previously (Jarrom *et al.*, 2009). In all experiments the ionic strength was kept constant, equivalent to 50 mM NaCl. The effect of addition of succinate at varying concentrations of NFZ is shown in Figure 5-17, and at varying NADPH concentrations in Figure 5-18. Succinate was a competitive inhibitor relative to NADPH (constant k_{catapp}, higher K_{mapp}) and an uncompetitive inhibitor relative to NFZ (lower k_{catapp}, lower K_{mapp}, constant k_{cat}/K_m). Global fits to Equation 2-7 for the NFZ titrations and Equation 2-8 for the NADPH titrations gave the parameters in Table 5-6 and Table 5-7 respectively.

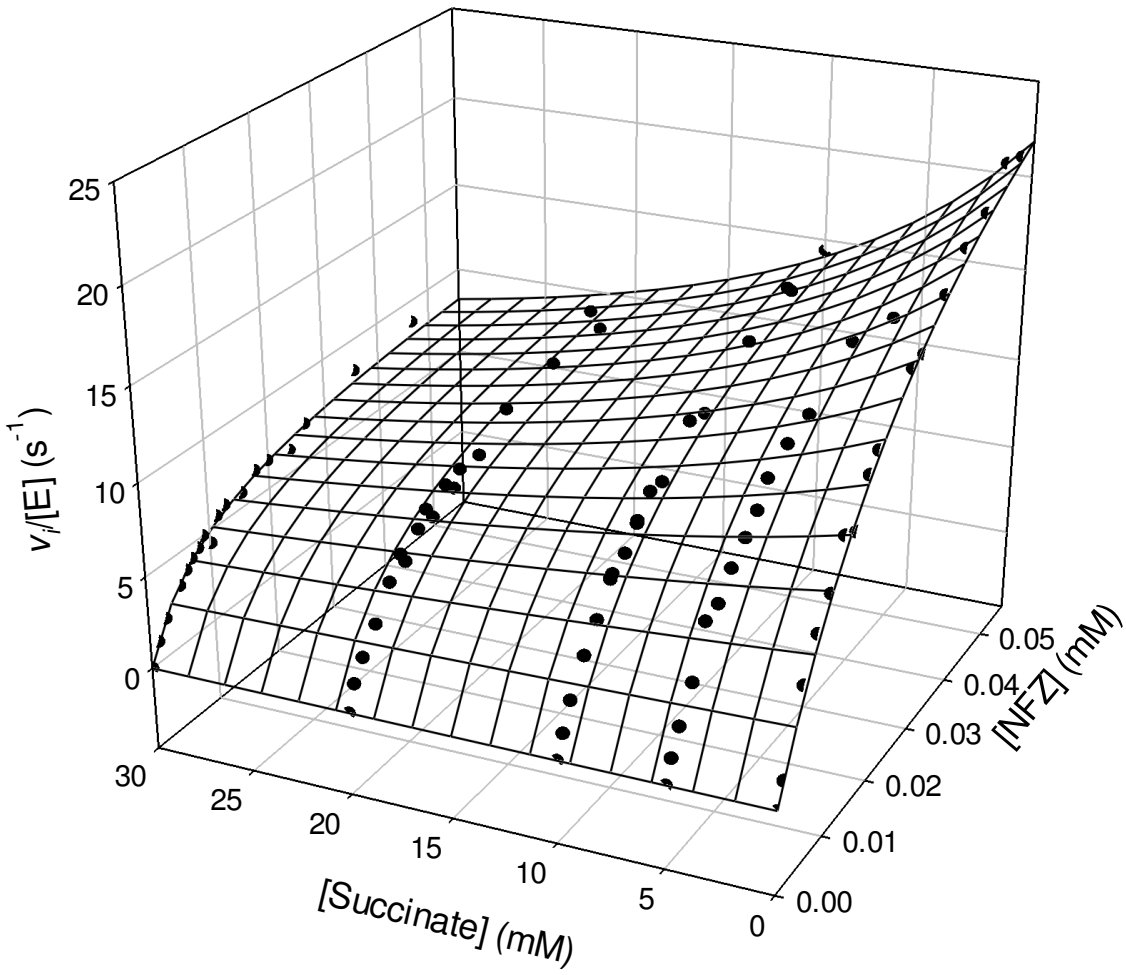


Figure 5-17: Global fit of data for succinate inhibition of the reduction of NFZ in the presence of NADPH by NfsA. NFZ titration data fitted to Equation 2-7**.** NFZ titrations were performed at 0, 5, 10, 20, and 30 mM succinate. All experiments performed at 100 μM NADPH.

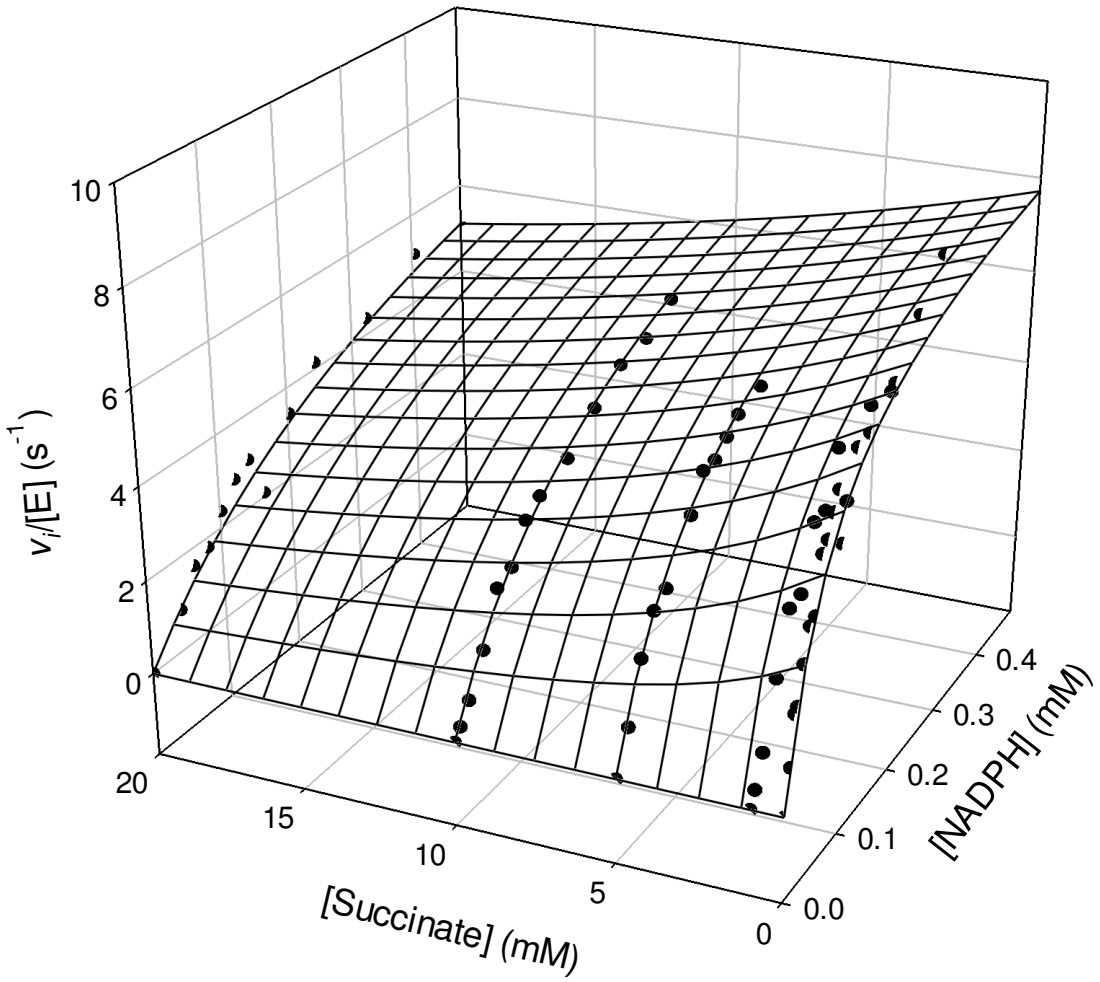


Figure 5-18: Global fit of data for succinate inhibition of the reduction of NFZ in the presence of NADPH by NfsA. NADPH titration data fitted to Equation 2-8**.** NADPH titrations were performed at 0, 1, 5, 10, and 20 mM succinate. All experiments performed at 100 µM NFZ.

[NFZ] µM	[NADPH] µM	[Succinate] mM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	K_{iU} mM	t	P
0 - 52	100	0 - 30	34.5 ± 1.5	23	<0.0001	32.9 ±2.7	12	<0.0001	11.1 ± 0.67	17	<0.0001

Table 5-6: Global kinetic parameters for the inhibition of the NfsA catalysed reduction of NFZ by sodium succinate at pH 7.0. Relative to NFZ, succinate shows uncompetitive inhibition and so the data was fitted to Equation 2-7 in order to calculate the K_{iU} inhibition constant.

[NFZ] µM	[NADPH] µM	[Succinate] mM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	K_{iC} mM	t	P
100	0 - 500	0 - 20	8.82 ± 0.44	20	<0.0001	68.8 ± 8.9	7.7	<0.0001	4.90 ± 0.54	9.0	<0.0001

Table 5-7: Global kinetic parameters for the inhibition of the NfsA catalysed reduction of NFZ by sodium succinate at pH 7.0. Relative to NADPH, succinate shows competitive inhibition and so the data was fitted to Equation 2-8 in order to calculate the K_{iC} inhibition constant.

5.11.2 Fumarate Inhibition

Similar experiments were performed with fumarate as an inhibitor. The effect of fumarate at varying concentrations of NFZ is shown in Figure 5-19, and varying NADPH in Figure 5-20. Fumarate was a competitive inhibitor relative to NADPH and showed mixed inhibition relative to NFZ. The global fit parameters are shown in Table 5-8 for the NFZ titration and Table 5-9 for the NADPH titration.

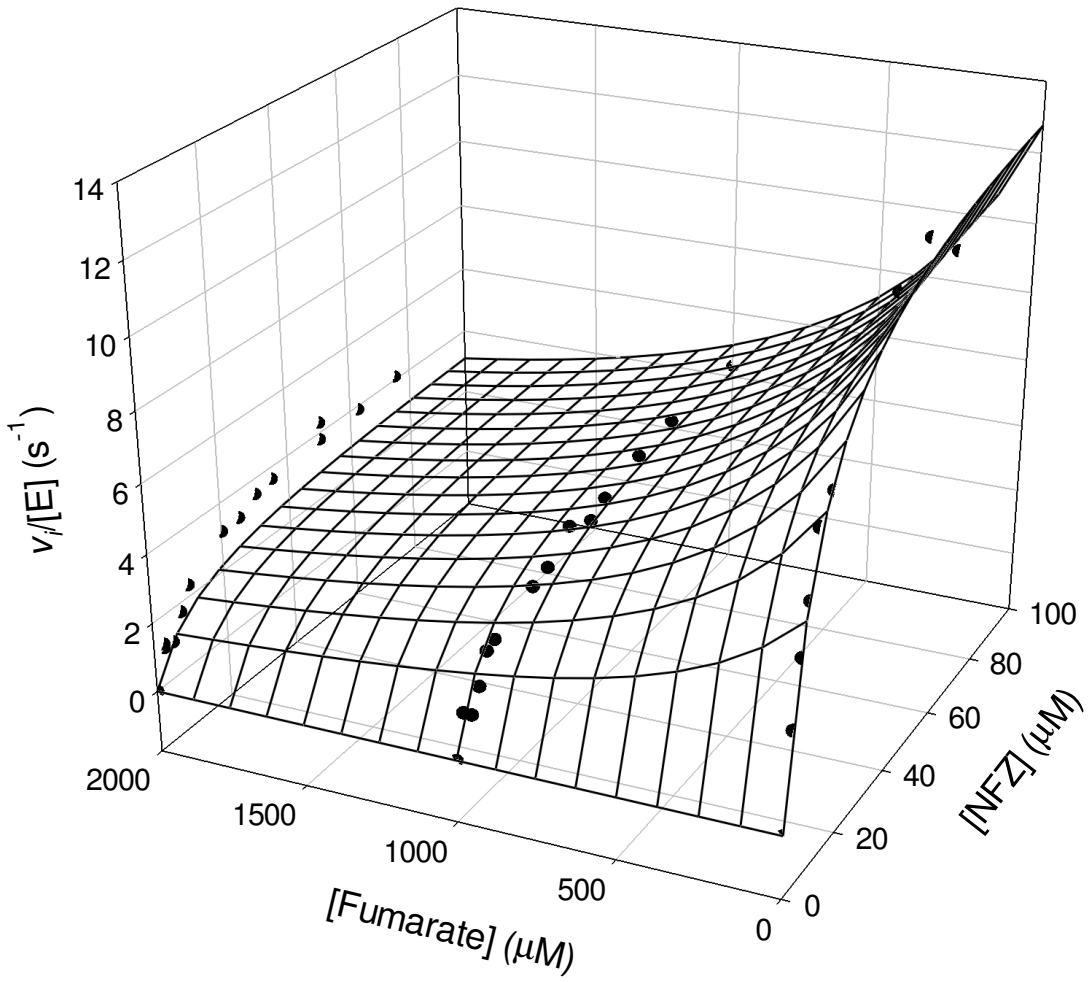


Figure 5-19: Global fit of data for fumarate inhibition of the reduction of NFZ in the presence of NADPH by NfsA. NFZ titration data fitted to Equation 2-9**.** NFZ titrations were performed at 0, 1 and 2 mM fumarate. All experiments performed at 100 μM NADPH.

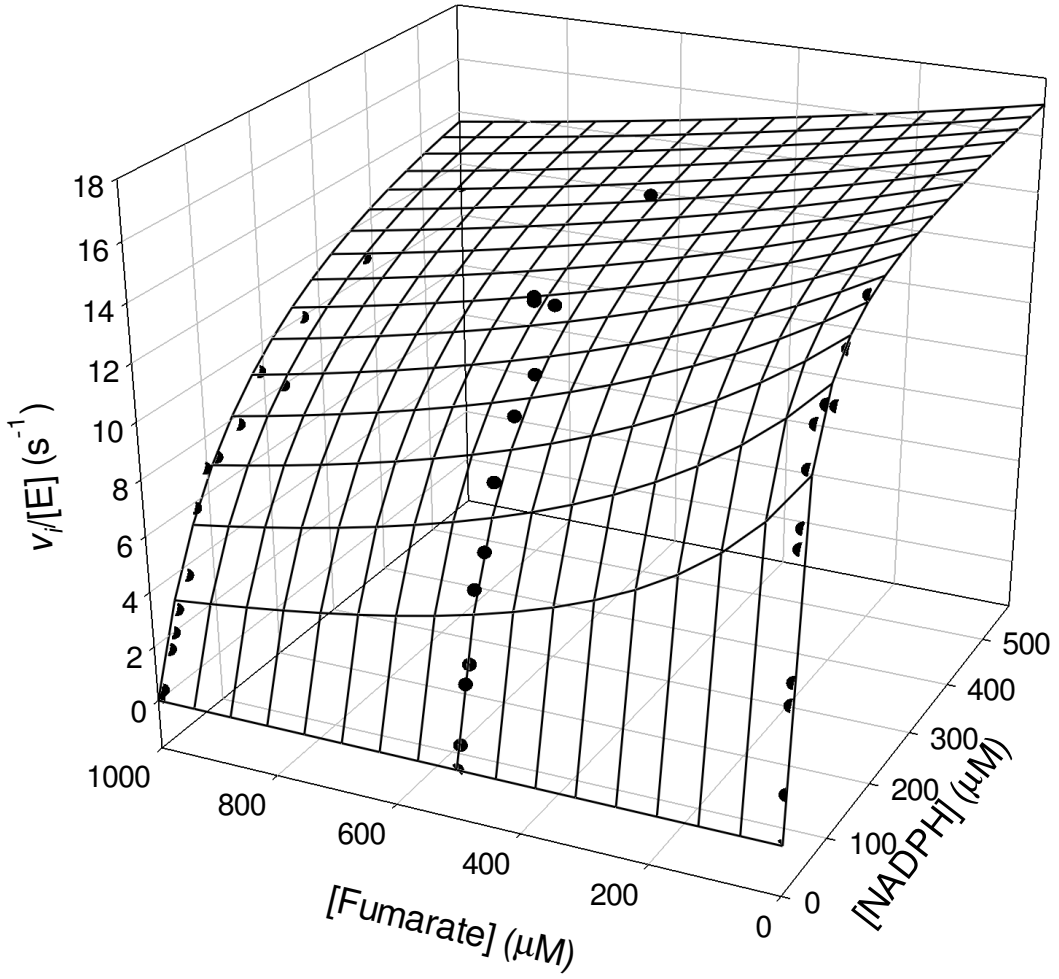


Figure 5-20: Global fit of data for fumarate inhibition of the reduction of NFZ in the presence of NADPH by NfsA. NADPH titration data fitted to Equation 2-8. NADPH titrations were performed at 0, 0.5 and 1 mM fumarate. All experiments performed at 75 μM NFZ.

[NFZ] µM	[NADPH] µM	[Fumarate] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	K_{iC} µM	t	P	K_{iU} µM	t	P
0 – 100	100	0 – 1000	14.1 ± 0.85	17	<0.0001	10.3 ± 1.9	5.4	<0.0001	546 ± 240	2.3	0.0267	653 ± 120	5.6	<0.0001

Table 5-8: Global kinetic parameters for the inhibition of the NfsA catalysed reduction of NFZ by sodium fumarate at pH 7.0. Relative to NFZ, fumarate shows mixed inhibition and so the data was fitted to Equation 2-9 in order to calculate the K_{iC} and K_{iU} inhibition constants.

[NFZ] µM	[NADPH] µM	[Fumarate] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	K_{iC} µM	t	P
75	0 – 580	0 - 2000	17.6 ± 0.85	21	<0.0001	18.1 ± 3.0	6.0	<0.0001	119 ± 17	6.9	<0.0001

Table 5-9: Global kinetic parameters for the inhibition of the NfsA catalysed reduction of NFZ by sodium fumarate at pH 7.0. Relative to NADPH, fumarate shows competitive inhibition and so the data was fitted to Equation 2-8 in order to calculate the K_{iC} competitive inhibition constant.

5.11.3 FMN Inhibition

Hana Ganan has performed some initial kinetic experiments investigating the inhibition of NfsA by FMN. Single titrations were performed of NFN (Figure 5-21) and NADPH (Figure 5-22), each with and without FMN. The estimates of the kinetic parameters are given in Table 5-10 and the errors and statistics are relatively poor due mostly to the low number of points. However it would appear that FMN is a competitive inhibitor with respect to NFN and uncompetitive with respect to NADPH. A rough estimate would be that the K_i was around 90 μM, making it quite a potent inhibitor.

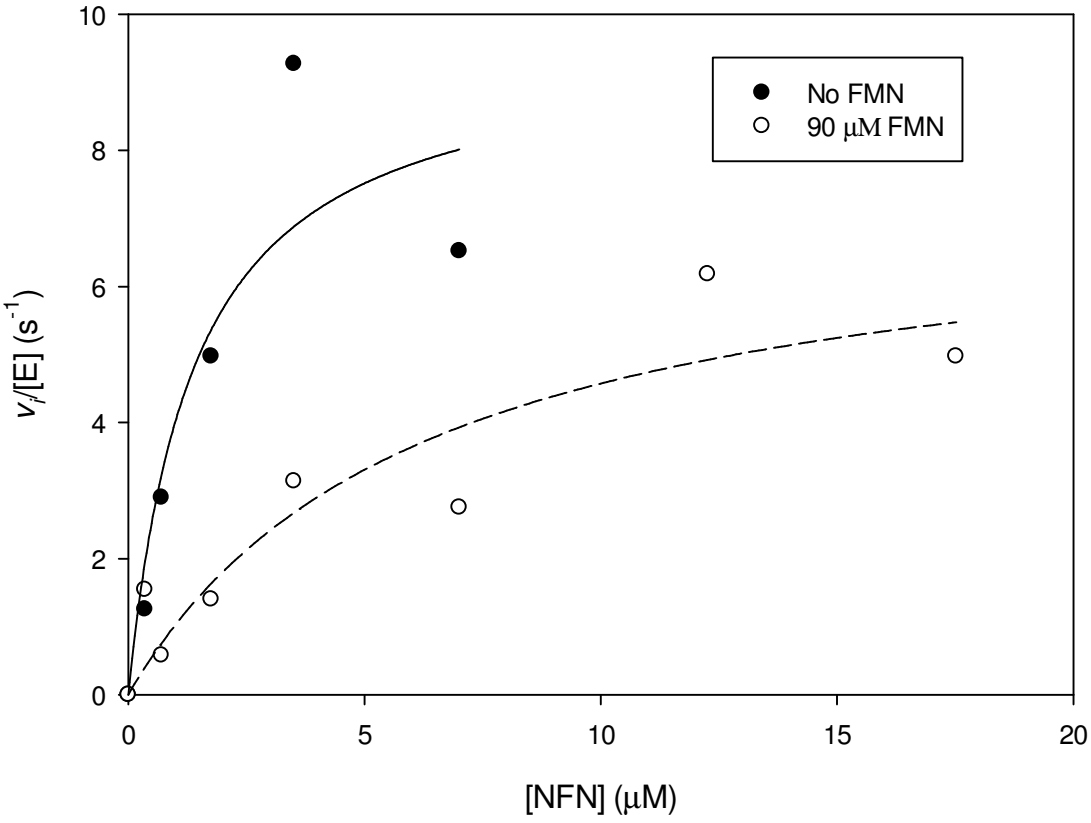


Figure 5-21: The effect of FMN on the rate of reduction of NFN by NfsA at 50 μM NADPH.

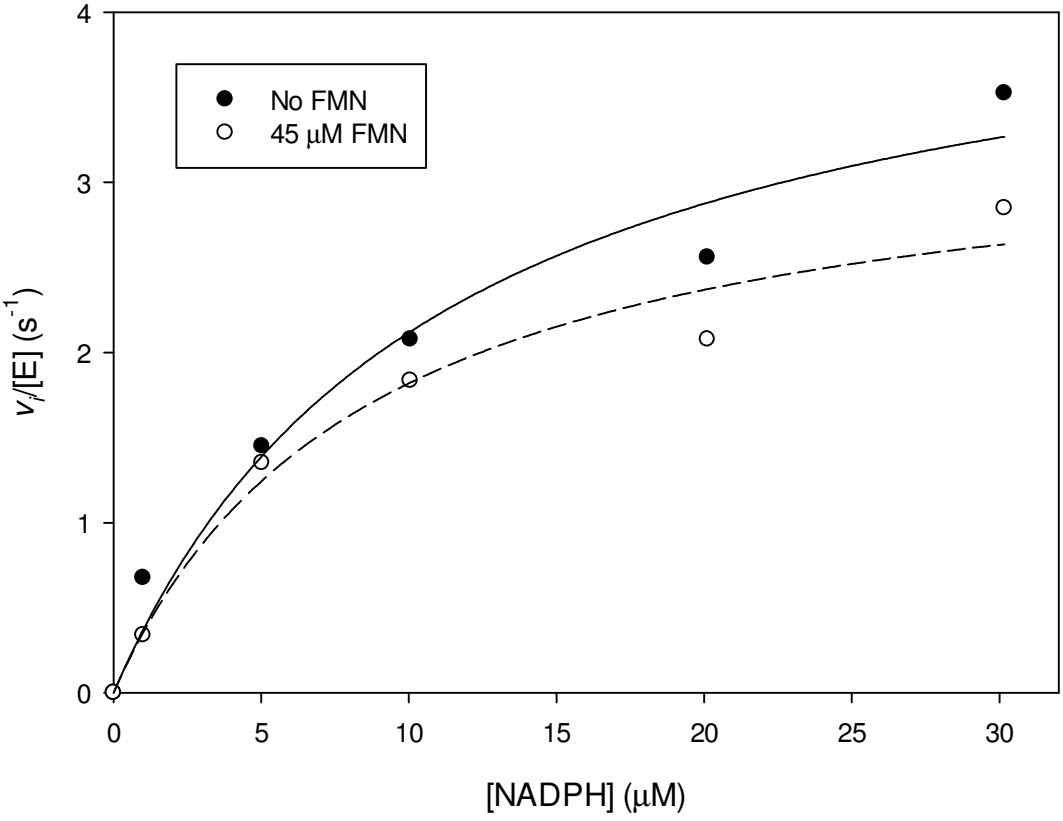


Figure 5-22: The effect of FMN on the rate of oxidation of NADPH by NfsA at 25 μM NFN.

[NADPH] µM	[NFN] µM	[FMN] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
50	0 – 17	0	9.60 ± 2.4	4.0	0.0162	1.38 ± 1.0	1.4	0.2427	6.94 ± 3.6	1.9	0.1288
50	0 – 17	90	7.42 ± 2.1	2.1	0.0120	6.22 ± 4.3	1.5	0.1970	1.19 ± 0.52	2.3	0.0626
0 - 30	25	0	4.48 ± 0.69	6.5	0.0029	11.1 ± 4.3	2.6	0.0595	0.402 ± 0.098	4.1	0.0148
0 – 30	25	45	3.39 ± 0.41	8.3	0.0011	8.66 ± 2.9	3.0	0.0398	0.392 ± 0.089	4.4	0.0115

Table 5-10: Steady-state kinetic parameters for the reduction of NFN by NfsA and NADPH in the presence and absence of FMN.

5.11.4 Quinone/Quinol Kinetics

Initial experiments performed by Hana Ganan attempted to analyse the kinetics of the reduction of quinone as catalysed by NfsA. Accurate K_m values proved difficult to measure as at concentrations of both NADPH and quinone that would give the lower part of the Michaelis-Menten curve the reaction was often completed within the time required to mix the solutions and begin monitoring the reaction. This leads to accurate estimates of k_{catapp}, but poorer estimates of K_{mapp} and k_{cat}/K_m. Attempts were made to couple the reaction to the reduction of cytochrome c as described for menadione in Section 2.2.3.3; however quinone reduces cytochrome c at a fast rate in the absence of the enzyme. The quinone titration is shown in Figure 5-23 and the NADPH titration is shown in Figure 5-24, and the kinetic parameters obtained are detailed in Table 5-11 and Table 5-12 respectively.

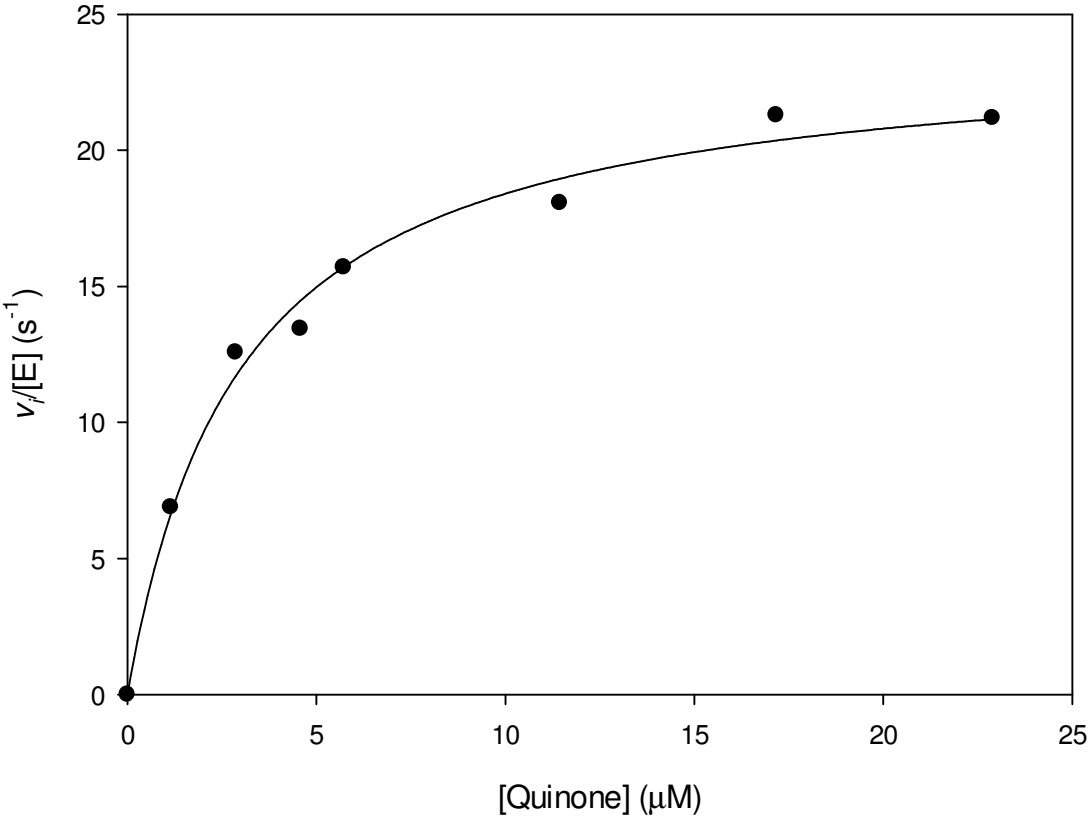


Figure 5-23: The initial rates of the reduction of quinone by NfsA in the presence of 50 μM NADPH.

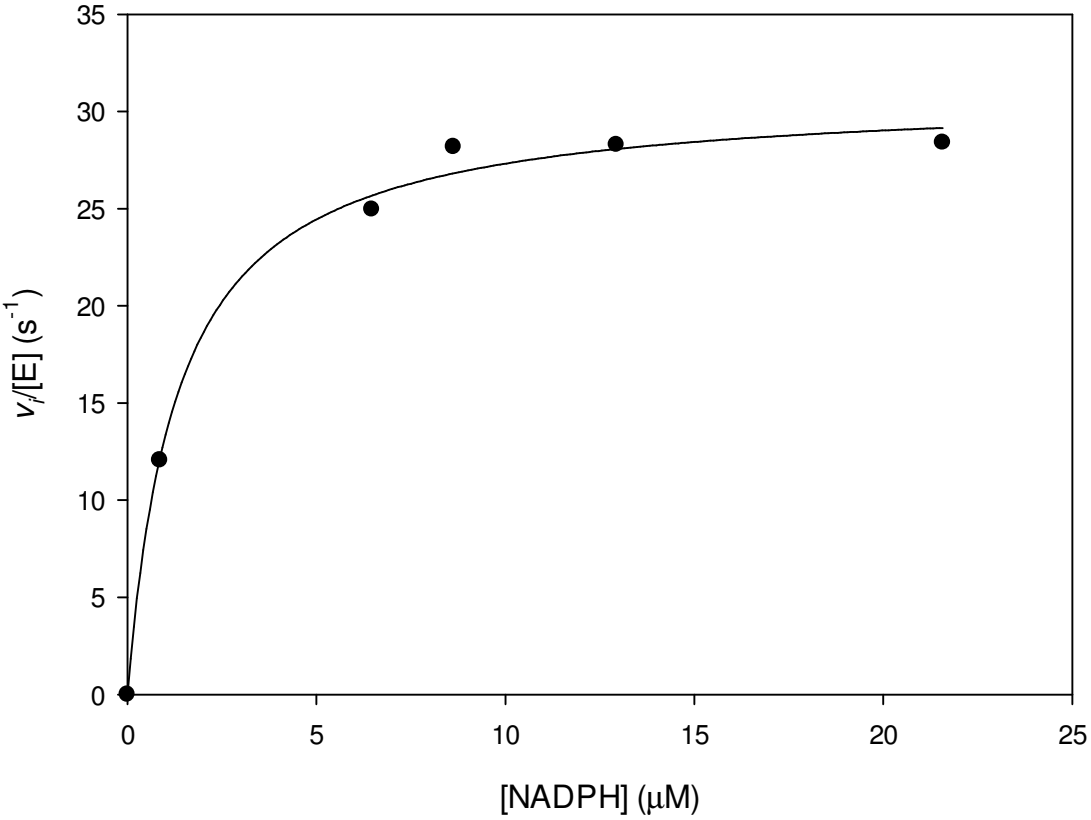


Figure 5-24: The initial rates of the reduction of quinone by NfsA in the presence of 50 μM quinone.

[NADPH] µM	[Quinone] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
50	0 - 23	23.9 ± 0.84	28	<0.0001	2.98 ± 0.37	8.1	0.0002	8.01 ± 0.76	11	<0.0001

Table 5-11: Kinetic parameters for the reduction of quinone by NfsA with NADPH. Data for the quinone titration at 50 µM NADPH.

[NADPH] µM	[Quinone] µM	k_{catapp} s^{-1}	t	P	K_{mapp} µM	t	P	k_{cat}/K_m $\mu M^{-1}s^{-1}$	t	P
0 - 22	50	30.9 ± 0.74	42	<0.0001	1.33 ± 0.18	7.3	0.0019	23.3 ± 2.8	8.4	0.0011

Table 5-12: Kinetics parameters for the oxidation of NADPH by NfsA in the presence of quinone. Data for the NADPH titration at 50 µM quinone.

5.11.5 Nitrofurantoin Kinetics

In obtaining global kinetic data for the reduction of NFN by NfsA and NADPH there were difficulties getting readings at high NFN concentrations (due to high background absorbance) and at low NADPH concentrations (due to the reaction being complete within the mixing time). The reaction was monitored at 400 nm and fitted to Equation 2-5 giving the global fit seen in Figure 5-25. The global kinetics values are given in Table 5-13.

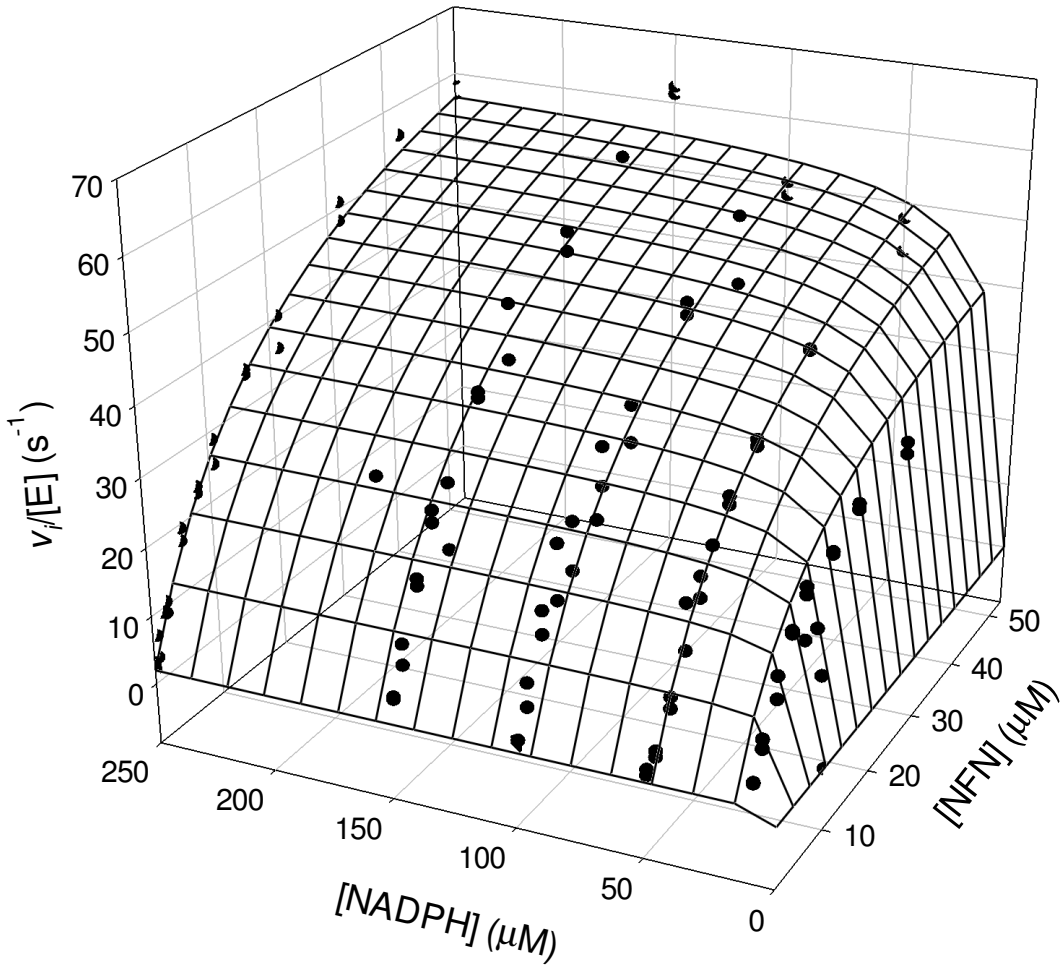


Figure 5-25: The effect of varying [NADPH] and [NFN] on the rate of reduction of NFN by NfsA at pH 7.0. Reaction monitored at 400nm.

		t	P
k_{cat} (s^{-1})	80.8 ± 3.1	26	<0.0001
K_m *NFN* (μM)	20.6 ± 1.6	13	<0.0001
K_m *NADPH* (μM)	10.9 ± 1.6	6.9	<0.0001
k_{cat}/K_m *NFN* (s^{-1} μM^{-1})	3.92 ± 0.18	22	<0.0001
k_{cat}/K_m *NADPH* (s^{-1} μM^{-1})	7.40 ± 0.97	7.6	<0.0001

Table 5-13: Global kinetic parameters for the reduction of NFN by NfsA and NADPH, monitored at 400 nm.

5.12 Stopped-Flow Kinetics

Figure 2-8 details the stages of the bi-bi ping-pong reaction catalysed by NfsA. Stopped-flow kinetics has been used to analyze the individual stages involved in this reduction as described in Section 2.6.2.

5.12.1 Oxidative Half-Reaction

In order for the reaction to be pseudo first-order, the concentration of substrate should be >2-fold higher than the concentration of enzyme. Since the reactions were followed by monitoring the reduction of the NfsA-bound FMN prosthetic group the concentration of enzyme had to be around 5 μM to give an observable change in

absorbance. However since the K_m for NADPH with NfsA is around 10 μM is was not possible to see the early part of the hyperbola binding curve and only the plateau was observed. A range of NADPH concentrations from 25 μM – 500 μM gave rates at ~ 160 s^{-1}.

5.12.2 Reductive Half-Reaction

For the reductive half-reaction, the reduction of CB1954, menadione and NFZ was analysed within an anaerobic chamber to prevent oxidation of the NfsA-FMN complex. There was a small degree of oxygen leakage causing slow oxidation, however data was fitted to the faster rates. For NFZ, two rates were seen at many concentrations, a quick rate (too quick to be measured at high NFZ concentration), and a slow rate seemingly independent of NFZ-cencentration. Thus it was not possible to get stopped-flow kinetic parameters for NFZ with NfsA.

5.12.2.1 CB1954 Reduction

Examples of the stopped-flow kinetic traces for each CB1954 concentration are shown in Figure 5-26. Each trace was fitted to a first-order reaction, Equation 2-17, and the rate constants were then plotted against CB1954 concentration in Figure 5-27. Only the early part of the kinetic curve from Figure 5-27 was measured, and above 500 μM CB1954 the reaction seemed to be complete within the deadtime. These rate constants were then fitted to Equation 2-18 and Equation 2-19 (Michaelis-Menten kinetics) to determine the rate constants and dissociation constants for the reductive half reaction, the stopped-flow kinetic parameters (Table 5-14).

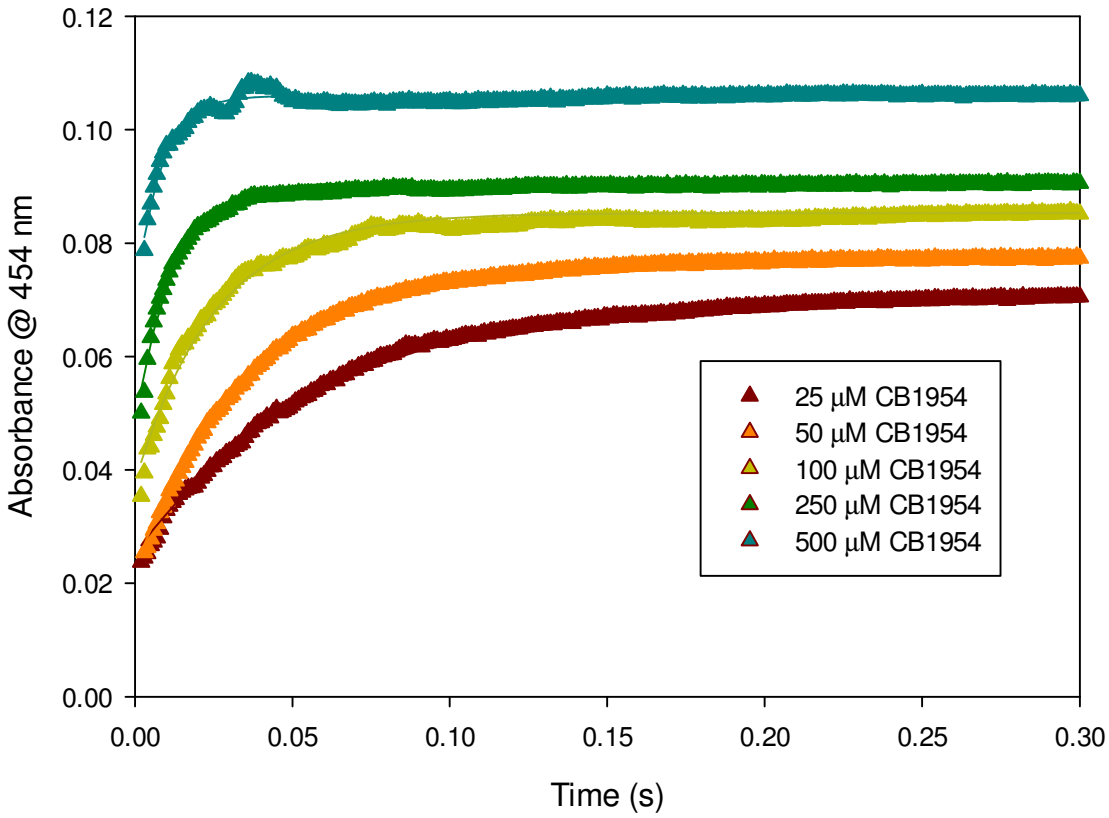


Figure 5-26: Stopped-flow kinetics for the reduction of CB1954 by reduced NfsA. Examples are shown for each concentration.

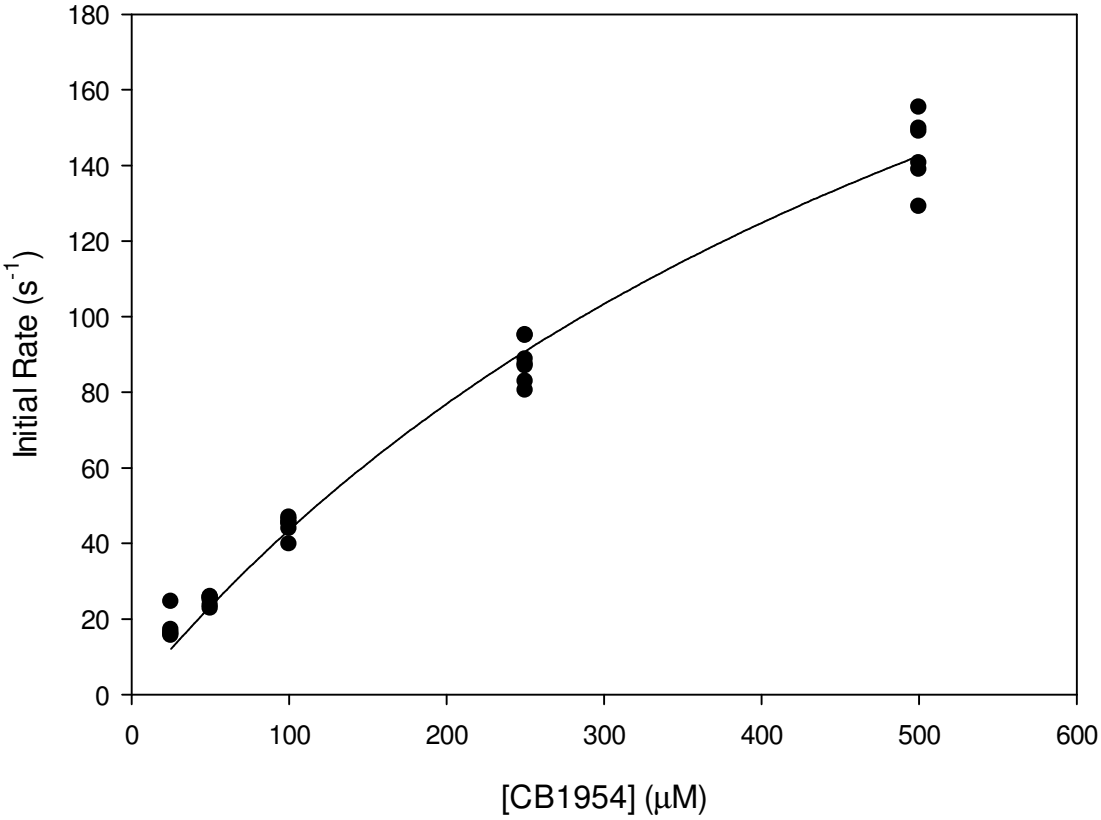


Figure 5-27: Plot of rate constants versus [CB1954] for the stopped-flow kinetic reduction of CB1954 by reduced NfsA. Data has been fitted to a hyperbola.

k_5 (s^{-1})	t	P	K_d (µM)	t	P	k_5/K_d (s^{-1} µM^{-1})	t	P	$k_{cat}/K_{mCB1954}$ (s^{-1} µM^{-1}) (Jarrom, 2008)
330 ± 28	12	<0.0001	659 ± 88	7.5	<0.0001	0.501 ± 0.025	20	<0.0001	0.23 ± 0.02

Table 5-14: Stopped-flow kinetic parameters for the reduction of CB1954 by reduced NfsA. The k_{cat}/K_m for CB1954 is taken from the Global kinetics in the reference (Jarrom, 2008).

5.12.2.2 Menadione Reduction

Examples of the stopped-flow kinetic traces for the reduction of menadione by reduced NfsA are shown in Figure 5-28. Each of these was fitted to Equation 2-17. The determined rate constants were plotted against menadione concentration (Figure 5-29). The fits of the kinetic traces were worse than for CB1954, as can be seen in Figure 5-28. As with CB1954, at high concentrations of substrate (>100 µM) the reaction was complete within the deadtime and so only the initial part of the curve is seen. This is reflected in the statistics in Table 5-15, where the estimates of k_5 and K_d are poor, but the estimate of k_5/ K_d is better.

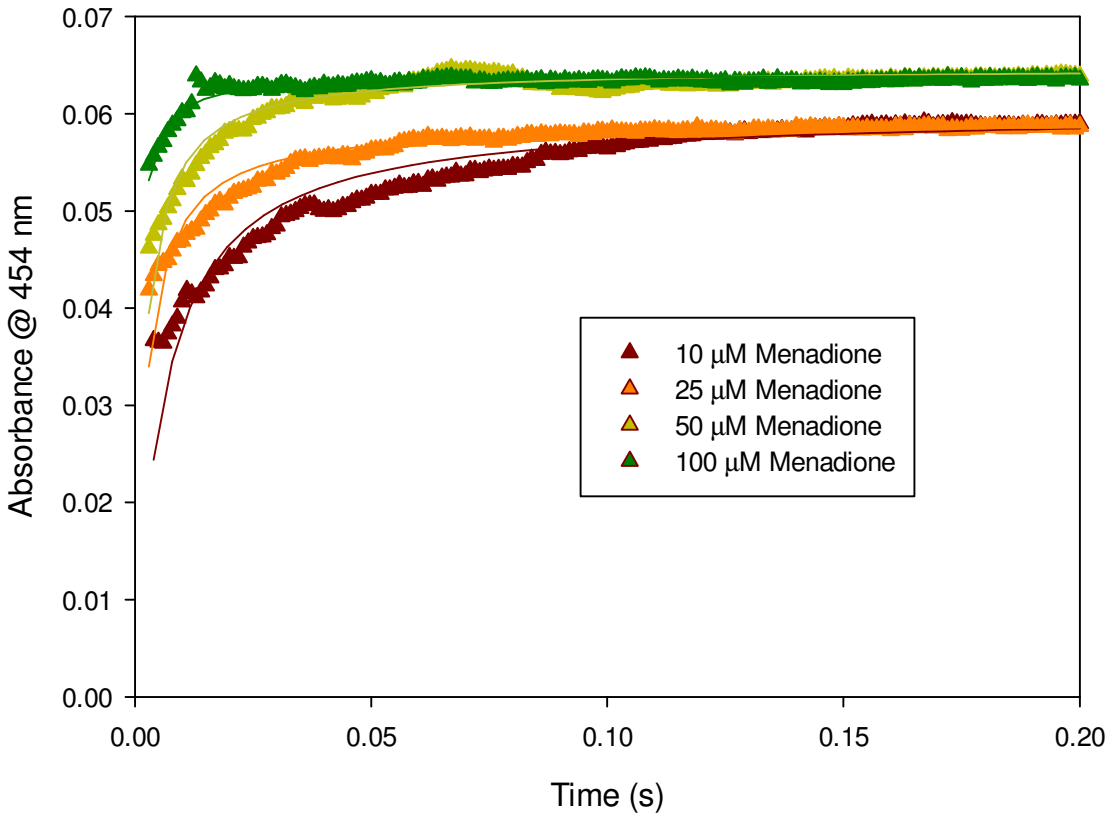


Figure 5-28: Stopped-flow kinetics for the reduction of menadione by reduced NfsA. Examples are shown for each concentration.

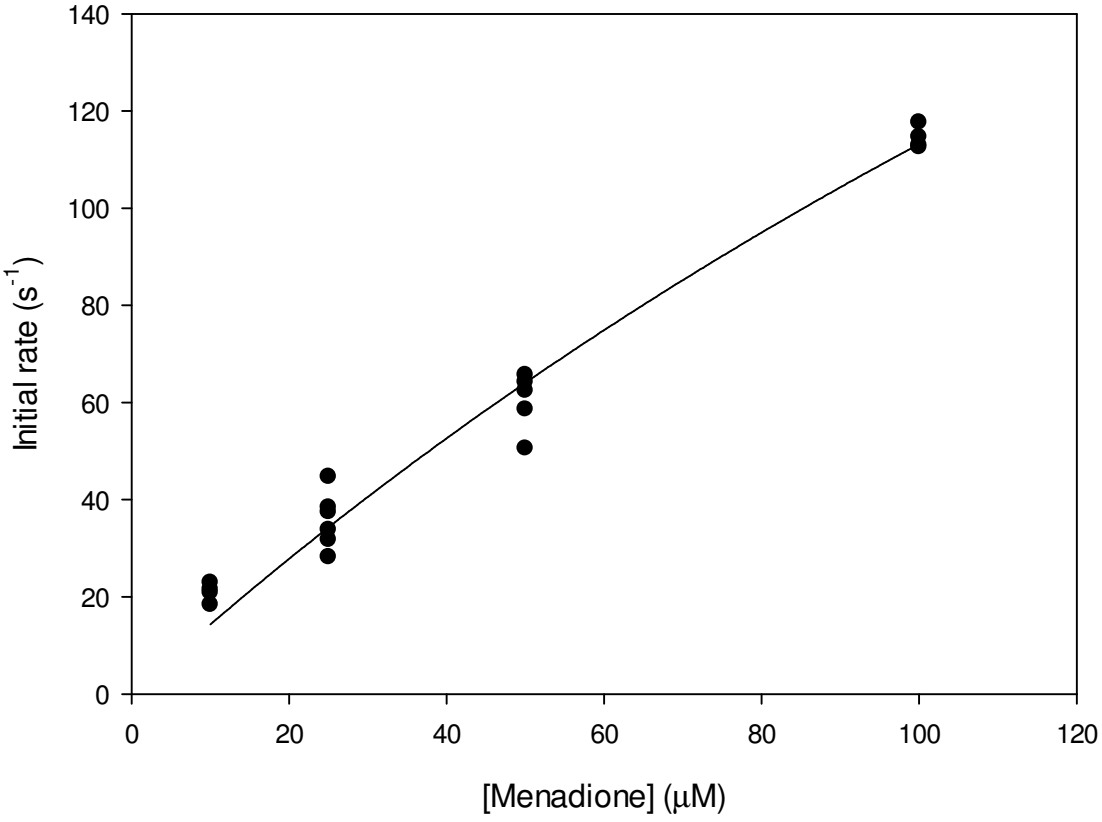


Figure 5-29: Plot of rate constants versus concentration of menadione for the stopped-flow kinetic reduction of menadione by reduced NfsA. Data has been fitted to a hyperbola.

k_5 (s^{-1})	t	P	K_d (μM)	t	P	k_5/K_d (s^{-1} μM^{-1})	t	P	$k_{cat}/K_{mmenadione}$ (s^{-1} μM^{-1}) (Jarrom, 2008)
485 ± 140	3.5	0.0027	328 ± 120	2.8	0.0112	1.48 ± 0.10	14	<0.0001	0.8 ± 0.07

Table 5-15: Stopped-flow kinetic parameters for the reduction of menadione by reduced NfsA.

5.13 Discussion

5.13.1 Analysis of Steady-State kinetics

The statistics presented for the steady-state inhibition by succinate and fumarate show that the K_i's for fumarate (0.5-0.6 μM) are ~10-fold lower than those for succinate. This increased binding strength with fumarate would be expected to be due to an addition π-π interaction of the central double bond in fumarate with the FMN aromatic ring system. The quinone reduction data is limited by the tight binding of the substrate to NfsA resulting in poor estimates of K_{mapp} but good estimates of k_{catapp}. The FMN inhibition data also presented relatively poor statistics, and further study is warranted. The global data presented here for NFN, shows it to have a >2-fold larger k_{cat} than is seen for either CB1954 or NFZ, and a higher K_m than NFZ, but lower than CB1954 (Jarrom, 2008).

5.13.2 Analysis of Stopped-Flow kinetics

Much of the stopped-flow data presented in this chapter showed some degree of fluctuation or a second rate. Also the maximum rate observed was not what would have been expected from the deadtime calculated for the machine in Bristol (see section 2.6.2.2), at around 160 s^{-1}. There were subsequently found to be several problems with both the machine and the anaerobic chamber (personal communication, Dr Ross Anderson), however the data presented here would still appear to be accurate as the other rates seen were distinct and much lower.

The plots of rate constants against substrate concentration for CB1954 (Figure 5-27) and menadione (Figure 5-29) do not show complete hyperbolas, only the initial roughly linear phase is seen. It was possible to fit the data to a hyperbola but the estimates of K_d and k_5 (for menadione especially) are relatively poor, low t values and high P values. The estimates of the initial slope of the curve, which gives k_5/K_d, are much better, allowing a comparison with previously determined k_{cat}/K_m values from steady state kinetics (Jarrom, 2008). The k_5/K_d values are 2-fold higher than the k_{cat}/K_m values for both CB1954 (Table 5-14) and menadione (Table 5-15). Even with the previously mentioned problems in observing fast rates of reaction, it is clear that for each of the substrates (including NADPH) that the stopped-flow rates are at least 4-fold higher than those seen in the steady-state (Table 5-16). This implies that, as was seen for WT NfsB (Jarrom *et al.*, 2009), the rate limiting step is likely to be product release.

	Max. Stopped-Flow rate (s^{-1})	**k_{cat} (s^{-1})**
NADPH	>160	-
CB1954	>160	42 ± 1
Menadione	>120	9.0 ± 0.4*
NFZ	>160	29.6 ± 0.9

Table 5-16: A comparison of the rates of NfsA catalysed reaction seen in the stopped-flow with the global steady-state rates previously determined (Jarrom, 2008). *Actually k_{catapp} taken from experiment at 100 μM NADPH.

5.13.3 FMN as a Potential Substrate

A single amino acid substitution, E99G, turns NfsA into a flavin reductase with a similar activity to FRP, the flavin oxidoreductase from *Vibrio harveyi* (Zenno *et al.*, 1998). This makes the structure of NfsA bound to FMN that of a pseudo-substrate. Figure 5-30 shows the position of Glu 99, and we can see that it H-bonds to Arg 133 and Arg 225. Arg 225 appears to be the most important residue in substrate binding, interacting with FMN, nitrofurantoin, and the inhibitors succinate and fumarate, whilst Arg 133 interacts with the prosthetic FMN. Glu 99 anchors these residues spatially and would have an electrostatic effect on the region of the protein. Mutating the Glu 99 residue to glycine would likely have two effects; firstly through the moving Arg 133, it would likely change the redox potential of the FMN. Secondly, it would remove the anchor of Arg 225 and allow this residue greater flexibility, possibly allowing the large 'substrate' FMN to completely enter the active site.

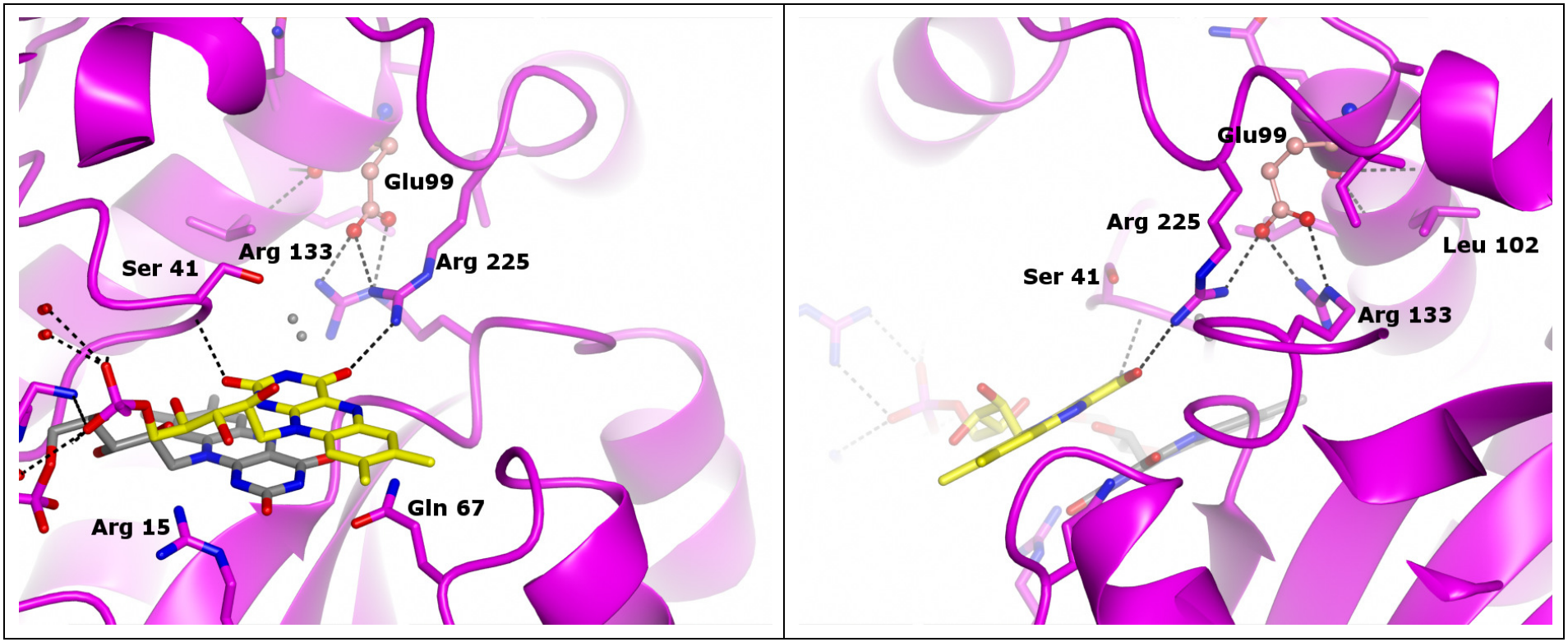


Figure 5-30: Location of the Glu 99 residue (pink) in the NfsA FMN structure. The FMN prosthetic group is shown in grey and the 'substrate' FMN is shown in yellow. The two images are 90 ° apart. Figures have been produced using CCP4 Molecular Graphics.

5.13.4 Clues to NADPH Binding

The change in the flexible loop in the NfsA FMN structure may hint at the mode of NADPH (structure in Figure 1-7) binding in NfsA, as the loop moves to form a phosphate binding pocket for the FMN phosphate. The mode of NADPH binding was investigated by Kobori *et al.* who proposed that the flexible loop (which had high B-factors) between helices I and J was involved in binding (for labelling see Figure 1-28). In their studies Arg 203 and Arg 208 were chosen for mutation as they appeared to be involved in NADPH binding in the NfsA analogue FRP (Tanner *et al.*, 1996). These residues are conserved across NfsA analogues, and while Arg 208 is facing away from the active site in all other structures of NfsA, it has been seen with other NADPH-binding enzymes that arginines can be moved to form phosphate-binding pockets (Wilson *et al.*, 1992, Reyes *et al.*, 1995, Adams *et al.*, 1994, Thorn *et al.*, 1995, Bolin *et al.*, 1982)]. The mutation R203A increased the K_m of NADPH ~33-fold, but had little effect on the K_m of NFZ, implying that Arg 203 is involved in NADPH binding, but the R208A mutation had little effect. Neither mutation had an effect on the K_m of NADH (although this is high) implying that Arg 203 interacts with the 2' phosphate of NADPH.

Kobori *et al.* proposed a binding mode for NADPH with NfsA based upon their experiments (Kobori *et al.*, 2001) and using the NADPH in the crystal structure of 6-phosphogluconate dehydrogenase (PDB I.D. 1PGO) (Adams *et al.*, 1994). This model has the nicotinamide moiety in a conformation opposite to that seen with nicotinate in NfsB (Lovering *et al.*, 2001), and the 2' phosphate of NADPH interacting with Asp 165, Tyr 199, Tyr 200 and Arg 203 forming a hydrophilic binding site. The NfsA

FMN structure presented here shows Arg 203 and Arg 208, along with Lys 167, Tyr 199 and Tyr 201 forming a binding-pocket for the flavin phosphate. This structure and the orientation of nicotinate in the NfsB active site have permitted production of another potential binding model of NADPH in NfsA (Figure 5-31). This model uses the major conformation of the NfsA FMN structure, and places the 2' phosphate into the density of the 'substrate' FMN phosphate. It differs from the model previously described in several ways. The change in the position of the flexible loop allows the NADPH to be much less stretched and the nicotinamide moiety to be placed in a conformation that is more amenable for hydride transfer (see section 1.6.1.1). The binding pocket for the 2' phosphate is much smaller and the ribose ring attached to the nicotinamide interacts with both the central and pyrollidinedione rings of FMN and Gln 67. Unlike the previous model, the adenine and the two central phosphates are out in solution. This would be logical given that this loop is present in NfsA but not in NfsB and given the preference of NfsA for NADPH. This implies that this loop should bind the 2' phosphate, with NfsA having extra interactions with phosphate in order to show a preference in cofactor. Effectively in this binding mode the binding would be primarily bidentate with the 2' phosphate bound in the phosphate binding pocket and the nicotinamide above the flavin rings.

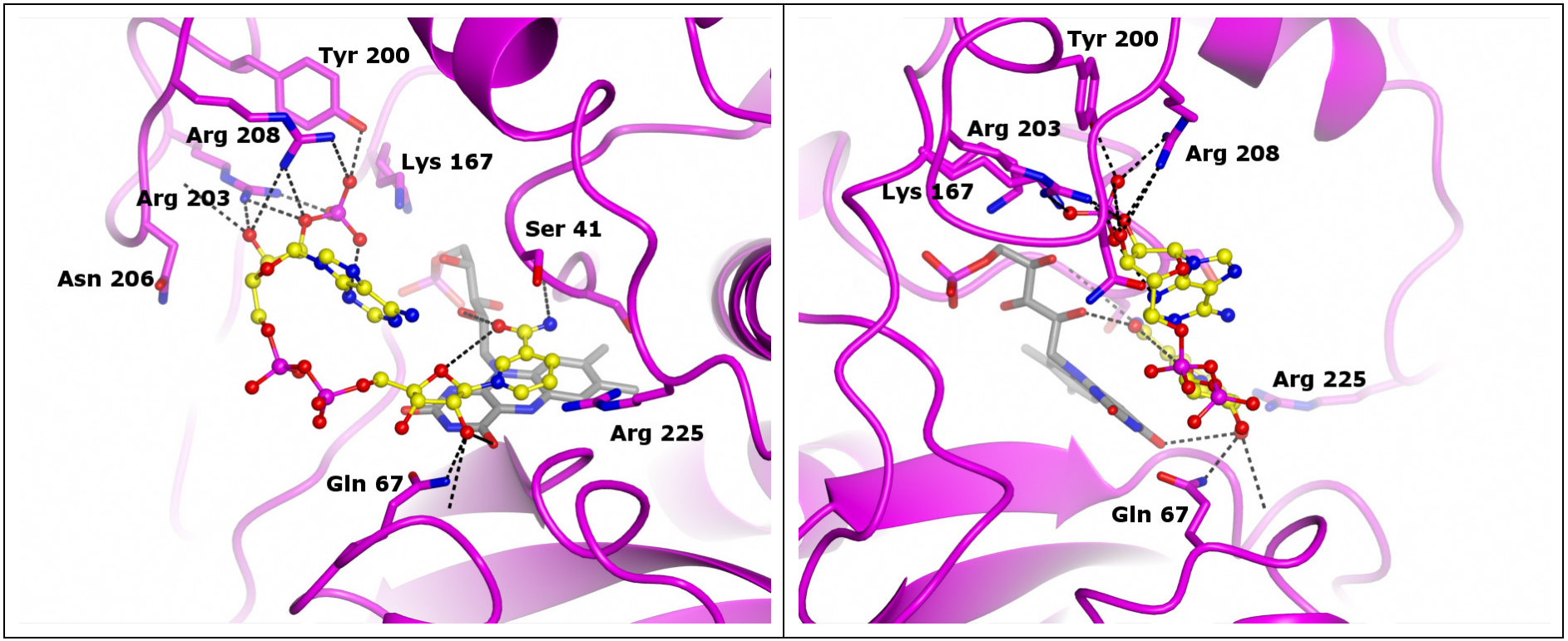


Figure 5-31: Proposed binding mode of NADPH in the active site of NfsA. The FMN prosthetic group is shown in grey and the NADPH in yellow. The protein is in the same conformation as the NfsA FMN structure, with the dominant loop conformation shown. The 2' phosphate of the NADPH has been modelled into the density of the FMN phosphate in the NfsA FMN structure. Two images are 90 ° apart. Figures produced using CCP4 Molecular Graphics.

With the loop between helices I and J being flexible, having been seen in two different conformations and always with the highest B-factors, the loop could move in a different way to accommodate the NADPH from that with FMN. It would seem likely that the movement would involve the formation of a phosphate binding-pocket involving Arg 203, and possibly Arg 208.

5.13.5 Modification of Cysteines

The NfsA FMN structure also contained modifications of two surface cysteines of NfsA, Cys 90 modified with the addition of a quinol (Figure 5-12) and Cys 9 modified with an imidazole modified quinol (Figure 5-14). The quinol-modified cysteine was also seen at Cys 9 in the NfsA Quinol structure, and density suggested another modification at Cys 9 but was inconclusive as to what was added in the NfsA Quinone structure. It is likely that further modification was not seen as the soaking times were decreased for each (10 min for NfsA FMN, 5 min for NfsA Quinol and 90 s for NfsA Quinone). Other structures including the NfsA Succinate 2 structure showed some modification of these cysteines, although always with density that was not clear enough to allow anything to be modelled. It is possible that something in the crystallization conditions somehow ‘activates’ these cysteines or that free radical damage from the X-rays causes modification of the residue. It is possible this is a method by which NfsA prevents oxidative damage.

5.13.6 Succinate/Fumarate as Clues to NfsA Evolution.

The structures of succinate and fumarate bound to the active site of NfsA may suggest an evolutionary history for the enzymes. The physiological role of the bacterial nitroreductases is unknown, and it may be that they have evolved from enzymes involved in redox reactions with dicarboxylic acids such as succinate and fumarate. The bond that is reduced in fumarate from succinate is in an 'active' position in the crystal structures presented here above the FMN N5 nitrogen.

5.13.7 NfsA Substrate Binding

The NfsA Quinone and NfsA Nitrofurantoin structures presented here are the first crystal structures of NfsA with known substrates bound in the active site. The most important residues in binding appear to be (i) Arg 225 through its guanidinium group, and (ii) the backbone of Ser41. Most other interactions of the substrate and other ligands seen here are with the FMN prosthetic group, although Arg 15 and Lys 167 also bind to some of the ligands. The crystal structures have given us insights into the important residues in the active site of NfsA, which residues are important for binding which substrates and why. This improves our understanding of the way NfsA binds its substrates.

6. Conclusions and Potential for Future Study

6.1 Summary of the Research Presented

Crystal structures have been determined of three NfsB natural amino acid mutants (Chapter 3) and two unnatural amino acid mutants (Chapter 4). These structures have allowed us to understand the chemical nature of the improvements in CB1954 reduction seen with these mutants. The T41Q/N71S/F124T mutant structure showed a hydrogen bond between the newly mutated Gln41 and Thr124 (Figure 3-8), and from the Gln41 to bound citrate, showing how the two mutations cooperate to bind ligands. The M127V mutation causes the smaller of two channels to the NfsB active site to open up (Figure 3-9), aiding access to the active site. The unnatural amino acids pNF and pAF, were both introduced in place of Phe124 and the crystal structures of both showed an effective rotamer change of the residue, placing the new polar group next to the substrate binding pocket (Figure 4-11 and Figure 4-14). The kinetics of the mutants have been further studied, and stopped-flow kinetics have shown that as the mutants have been selected for their CB1954 binding, NADH binding has suffered and should this trend continue then the rate determining step (currently product release), may become the rate of NADH oxidation (Section 3.13).

The crystal structure of NfsA has also been determined bound to the inhibitors succinate and fumarate, the substrates quinone (and product quinol) and nitrofurantoin, as well as a second bound FMN (Chapter 5). These structures have shown for the first time the residues involved in substrate binding, and the FMN bound structure has allowed a model for NADPH binding to be proposed (Figure

5-31). Steady-state kinetics of substrates and inhibitors found in the crystals have also been shown, showing that fumarate is a better inhibitor than succinate.

6.2 Comparison of NfsA and NfsB

6.2.1 Structural Comparison

NfsA and NfsB exhibit little sequential homology (16.1% identity, 27.2% similarity) but do exhibit some structural homology. Figure 6-1 shows a structural alignment of the crystal structures of NfsA and NfsB dimers, whilst Figure 6-2 shows a sequence alignment with structural elements shown. Both enzymes have a 5-stranded anti-parallel β-sheet core surrounded by α-helices. The active site is located at the dimer interface with the *re*-face of the isoalloxazine ring of the FMN presented towards the substrate pocket. The areas shown within the black circles on Figure 6-1 have each structural element matched up to one from the other enzyme, although by no means perfectly aligned. It is the outer helices where the structure of the enzymes differ the most, especially helices I and J in NfsA and E and F in NfsB. These differences may help to explain NfsA's preference for NADPH over NADH, as it is the loop between helices I and J that has been shown to be involved in NADPH phosphate binding (Kobori *et al.*, 2001). The sequence alignment shows that the majority of the conserved regions are early in the sequence where the structural elements also overlay well, however beyond the 100th residues the similarities in sequence are not matched by the structural elements.

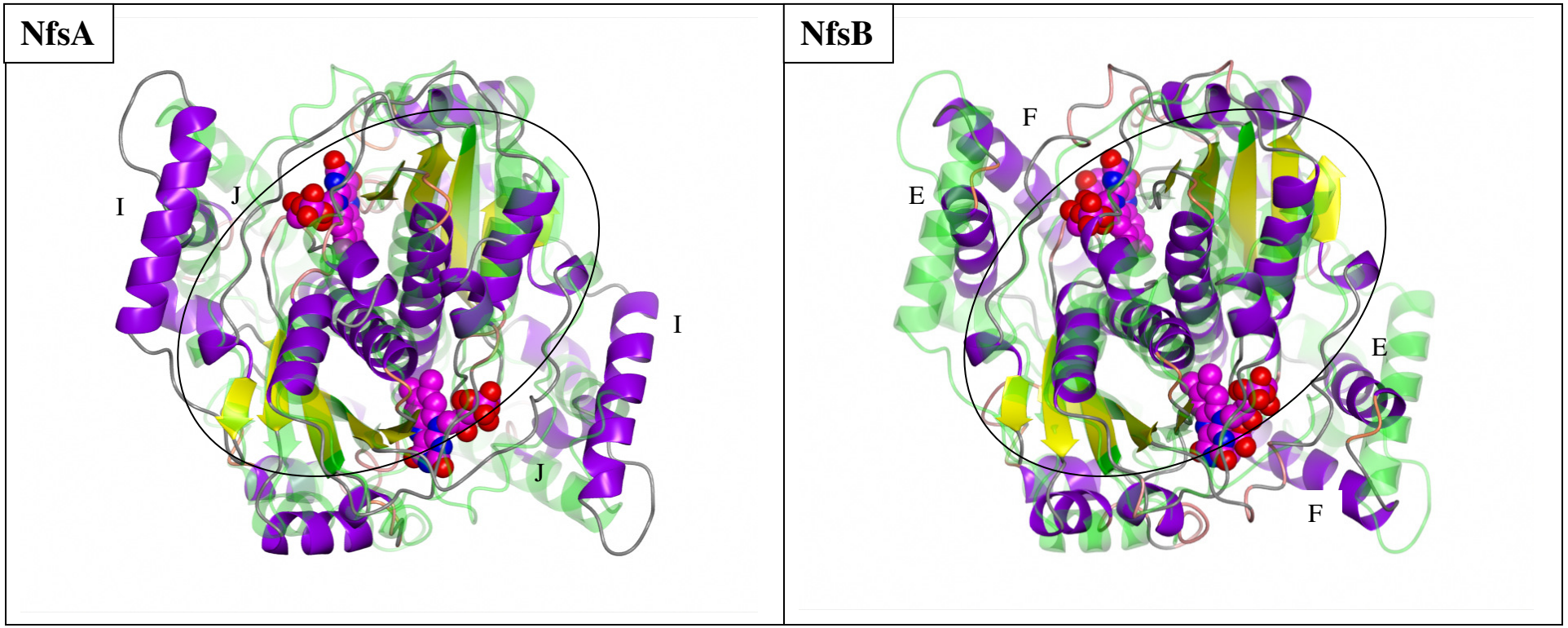


Figure 6-1: Comparison of the crystal structures of NfsA (right) and NfsB (left) shown with α-helices in purple and β-sheets in yellow. The two FMN prosthetic groups are shown as magenta spheres. The structures were aligned using THESEUS (Theobald and Wuttke, 2008). The black circles over each image contain the area of the protein where each structure element approximately aligns. In each image the other enzyme is shown as a transparent green. Figure produced using CCP4 Molecular Graphics.

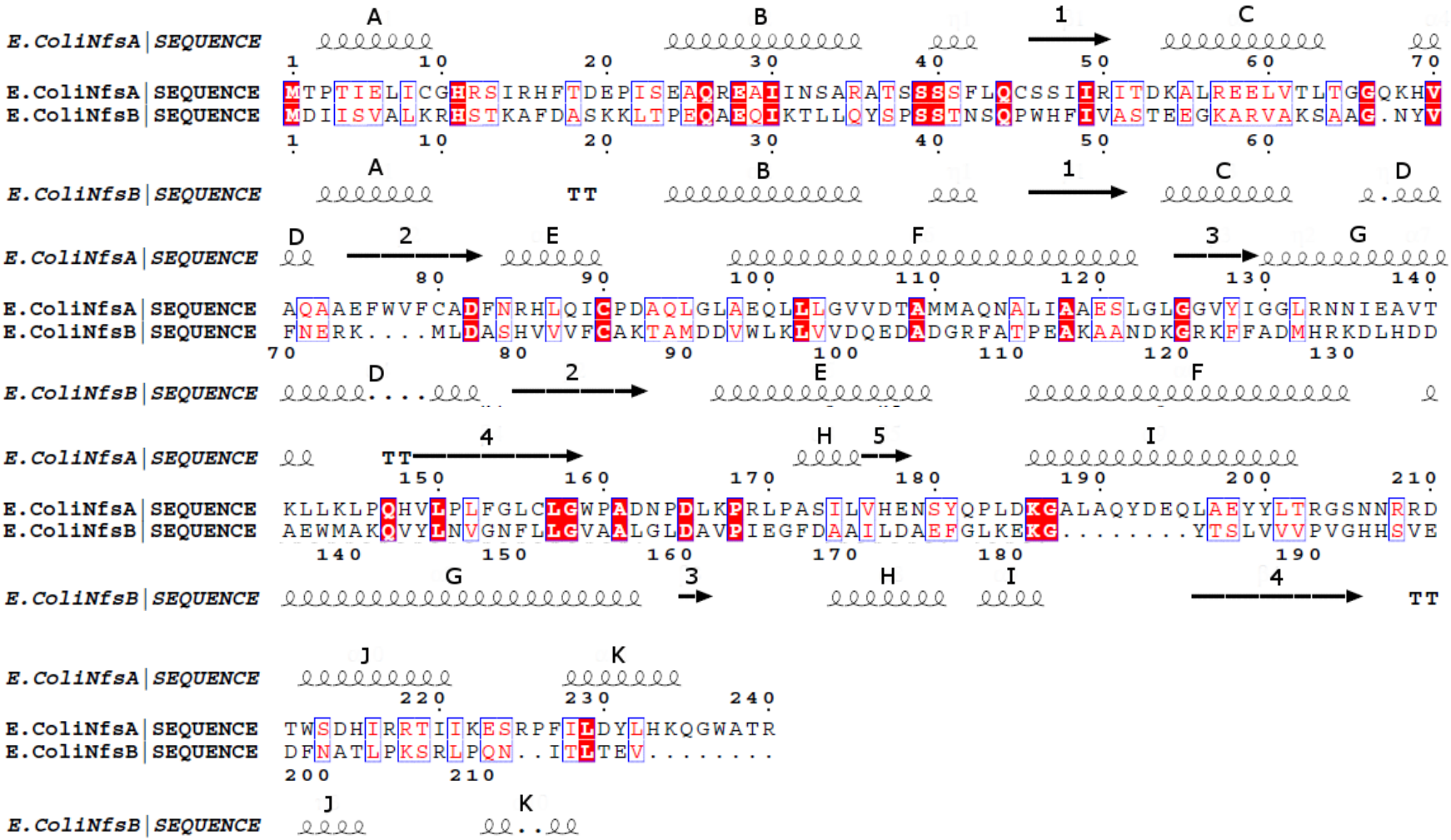


Figure 6-2: Sequence alignment of NfsA and NfsB showing secondary structural elements. Alignment performed using ClustalW2 (Larkin *et al.*, 2007), figure produced using ESPript (Gouet *et al.*, 1999) and subsequently edited to combined both secondary structural elements.

The active sites of NfsA and NfsB are compared in Figure 6-3, showing a numbers of functionally conserved residues shown in purple. Most of the residues directly interacting with the FMN prosthetic group are conserved in function across the two sites, although a few non-conserved interactions are seen. These interactions are summarised in Table 6-1. There are three conserved loops, the first roughly from residues 10-15 in each enzyme, which contains polar groups that primarily interact with the flavin phosphate. The residues here are shifted by one amino acid between the proteins, but the nature of the hydrogen bonding interactions are conserved for the δ^+ or δ^- charge of the hydrogen bonding atom. The second conserved loop, is from 39-41 in both enzymes, and is conserved for side chain function, although most of the side chains point away from the active site. The third loop (residues 128-131 in NfsA and 163-166 in NfsB) interacts with the *re* face of the FMN. This is mostly a backbone interaction and so the only conserved residue is the final of these 4, which is a glycine in both enzymes.

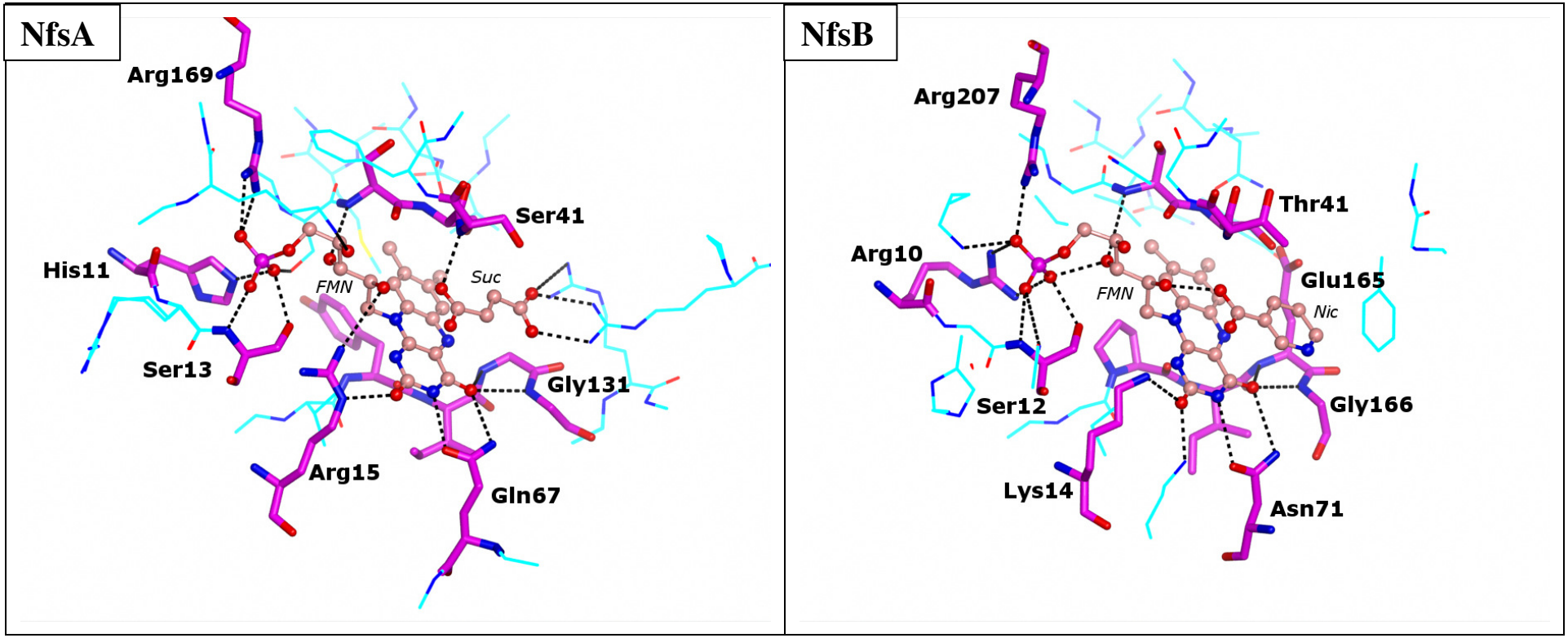


Figure 6-3: A comparison of the active sites of NfsA (left) and NfsB (right). The NfsA Succinate 1 structure is compared with the WT-NfsB nicotinate-bound structure (pdb I.D. 1ICR (Lovering *et al.*, 2001)) as the ligands are a similar size. All residues within 4.0 Å of either the FMN or the succinate/nicotinate, which are themselves shown in pink. Of the surrounding residues, those with conserved function between the two enzymes are shown in purple, the rest are shown in blue. Some key residues are labelled. Figures produced with CCP4 Molecular graphics.

Residue in NfsA	Residue in NfsB	Function	Side chain or main chain interaction
His11	Arg10	Hydrogen bonding to FMN phosphate	Side
Ser13	Ser12	Hydrogen bonding to FMN phosphate	Side
Arg15	Lys14	Hydrogen bonding to FMN ketone oxygen 'O2'	Side
Ser39	Ser39	Hydrophobic interactions with FMN	Main
Ser40	Ser40	Hydrophobic interactions with FMN	Main
Ser41	Thr41	Hydrophobic interactions with FMN and hydrogen bonding to substrate	Side and Main
Gln67	Asn71	Hydrogen bonding to FMN ketone 'O4' and nitrogen 'N3'	Side
128-131	163-166	Hydrogen bonding and hydrophobic interactions to FMN isoalloxazine ring system	Main
Arg169	Arg207	Hydrogen bonding to FMN phosphate	Side

Table 6-1: Functionally conserved interactions in the active sites of NfsA and NfsB.

6.2.2 Comparison for use in Gene Therapy

The best indication for the potential of each enzyme for use in GDEPT with CB1954 or other prodrugs will be the ability of the combination of prodrug and enzyme to shrink tumours and kill tumour cells. However, understanding the interactions in the active site of the enzymes can allow rational design of future drugs and drive further mutagenesis. NfsA has shown greater efficacy in GDEPT over NfsB *in vitro* than *in*

vivo (see Section 1.5.3.2). This difference between the *in vivo* and *in vitro* effect could be explained either by increased inhibition by small molecules in the cell or by relative levels of cofactors available. More cofactor molecules will be available to NfsB since it can utilize either NADH or NADPH. It has been estimated that the cellular concentration of NAD^+ and NADH is ~10-fold higher than the concentration of $NADP^+$ and NADPH (Reiss *et al.*, 1984), with cytosol concentrations (of NAD^+ and NADH) estimated at around 0.3 mM in animals (Yamada *et al.*, 2006, Yang *et al.*, 2007) and 1.0 – 2.0 mM in yeast (Belenky *et al.*, 2007). The two enzymes are inhibited by different molecules, for example FMN inhibits NfsA but not NfsB, whilst dicoumarol inhibits NfsB but not NfsA (Zenno *et al.*, 1996a). This study has shown that NfsA is inhibited by succinate and fumarate; both molecules are members of the Citric Acid Cycle and so will be present in most cells. Experiments not detailed here showed that neither inhibit NfsB. Fumarate levels in some human cells have been shown to be 20-40 μM (Koivunen *et al.*, 2007), and levels of succinate would presumably be similar. So the K_i's of succinate and fumarate and 100- and 10-fold higher than the estimated concentrations, implying succinate is unlikely to cause significant inhibition *in vivo* but fumarate will likely cause some low levels.

6.3 Rationalisation of Improvements Seen with NfsB Mutants

All of the NfsB mutants in this study reduce CB1954 with higher specificity constants than WT, and crystal structures can give an insight into how these improvements work. The effect of the T41L and N71S mutations have been discussed elsewhere for the crystal structures of the single mutants (Race *et al.*, 2007), and no additional changes are seen when the mutations are combined. The effect of removing the bulky

F124 has also been suggested to be primarily a steric effect, and mutants showed either multiple conformations of side chains or polar side chains moving away from the active site and into a polar pocket. The combination of T41Q and F124T provided more improved kinetics than expected from the combination of single mutations (Guise *et al.*, 2007, Jarrom *et al.*, 2009). The structures of the T41Q/N71S/F124T triple mutant and the T41Q/N71S/F124T/M127V quadruple mutant present here suggest a reason for the observed cooperativity. The two side chains interact with each other in the triple and quadruple mutants, through a hydrogen-bond between the glutamine NE2 and the threonine OG1 (Figure 3-11). These interactions may be important in the binding of larger substrates as can be seen when the inhibitor citrate is bound in the active site, as Glu 41 hydrogen bonds to the citrate as well as Thr 124. This may explain the synergy of the two mutations, through the increased interactions with substrates stabilised through the interaction between the side chains. The T41Q mutant was detrimental on its own, and this may be because without the F124T mutation the glutamine side chain may not be in the same position and may interfere with the action of the enzyme.

There are two channels leading into the active site, both shown in Figure 3-9 a). The smaller of the two is too small for most substrates but may be a path for solvent or protons. The M127V mutation opens up this smaller channel and effectively merges the two together (Figure 3-9 b)). In part it does this by replacing M127 with a smaller residue, but there are other changes induced including movement in the position of the ASN67 side chain. This would provide improved access for substrates and solvent to the active site. The T41Q/N71S/F124T/M127V quadruple mutant was selected by a

similar method to the T41Q/N71S/F124T triple mutant (described in Section 1.6.2.3.1), giving a similar improvement over WT in terms of IC50 for CB1954 (unpublished work by Simon Vass). However subsequent studies by undergraduate project student Alex Gavin (unpublished), suggest that the quadruple mutant binds CB1954 less tightly *in vitro* than the triple, but binds menadione more strongly than any previously observed mutant. So this opening up of the active site may be more beneficial for the larger menadione.

6.4 Clues to cofactor and substrate binding in NfsA

In Section 5.13.1 FMN was identified as a pseudo-substrate of NfsA, given a single amino acid substitution (E99G) converts NfsA to a flavin reductase with similar activity to Frp, the flavin reductase from *Vibrio harveyi* (Zenno *et al.*, 1998). Similarly NfsB can be converted to a flavin reductase with a mutation of Phe 124 to 10 other amino acids (Zenno *et al.*, 1996b), and its activity is similar to FRase I, the flavin reductase from *Vibrio fischeri*. Noticeably only one of these mutations (F124H) improved activity of NfsB with either NFZ or CB1954 (Grove *et al.*, 2003). Frp and NfsA share much sequence homology, as do FRase I and NfsB. The NfsA FMN structure may give clues to NADPH binding and the specificity of NfsA for the cofactor as it shows how NfsA can form a phosphate binding pocket (see Section 5.13.4). If FMN could be bound to NfsB it is possible that conclusions could be drawn about the presence or lack of a phosphate-binding pocket, however, if as speculated here, it is the phosphate binding pocket of NfsA that leads to the specificity, then it is likely that getting a structure of NfsB bound to FMN would prove difficult as it would lack this interaction.

6.5 Clues to mechanism of reaction

Section 1.6.1.1 details the possible mechanisms of oxidation and reduction in the FMN-bound nitroreductases. The reduction of the flavin by NAD(P)H and the reduction of quinones is likely to be by hydride transfer (simultaneous transfer of hydrogen and electrons from the FMN) as both can only be involved in two electron reductions. However a recent study has suggested that the reduction of CB1954 (and presumably other nitroaromatics) by NfsB happens by electron transfer from the flavin followed by proton transfer from solution (Christofferson and Wilkie, 2009) (discussed in Section 1.6.1.1). There was further evidence for this mechanism as the crystal structure of NfsB bound to NFZ (a nitroaromatic substrate) showed NFZ bound with the nitro group away from the reactive N5 nitrogen of the FMN (Race *et al.*, 2005) (Figure 1-22). If the mechanism were hydride transfer then NFZ in this structure would not be in an active conformation, however for electron transfer (transfer of two electrons from the flavin and one proton from solution) a fully conjugated substrate (such as NFZ and CB1954) would not require the nitro group to be above the N5 nitrogen. The NfsA Nitrofurantoin structure presented in Section 5.10 has the NFN bound with the nitro group away from the reactive centre of the flavin ring system. It is likely the chemistry of the active site of NfsA and NfsB have the same mechanism, so this adds to the evidence that the reduction of nitroaromatics by the *E. coli* nitroreductases occurs by an electron transfer mechanism.

There are other potential explanations for the nitro groups being bound away from the active site in the two crystal structures. Since the enzymes in the crystals are oxidised and the nitroaromatic substrates interact with the reduced enzyme complex, the binding is not necessarily representative of an active complex. It is possible that the change in the redox state of the FMN would lead to the substrates binding in a different orientation to that observed with the oxidised enzyme (as proposed previously Race *et al.* (Race *et al.*, 2005)). Another possibility is that the major binding orientation of the substrate in the enzyme is not the reactive conformation, which only occurs rarely. This has been proposed to be the case with pentaerythritol tetranitrate reductase based upon similar results (Khan *et al.*, 2005). However the combination of the two crystal structures and the biophysical calculations (Christofferson and Wilkie, 2009) suggests strongly that the reason for the unexpected orientation is that the mechanism is electron transfer.

6.6 Implications of Electron Transfer on New Prodrug Design

If the mechanism of reduction is electron transfer this would affect the rational design of further prodrugs based upon the crystal structures of the enzymes. Firstly the prodrug would not have to be orientated with the group to be reduced above the FMN N5 nitrogen. Secondly the drug would not have to be as close to the flavin ring system, although the closer the drug the faster the rate of electron transfer (see Section 1.6.1.1). This would increase the range of possibilities for prodrug design; the drugs could be designed to interact with any part of the active site, and in-particular for NfsA the important Arg225 residue.

6.7 Residues in NfsA to Target for Mutagenesis

The determination of the ligand-bound crystal structure of NfsB led to the identification of 9 residues to mutate and has since led to the discovery of mutants showing much improved activity with CB1954 (see Section 1.5.3.1). The determination of ligand-bound structures of NfsA (Chapter 5) would allow a similar targeting of residues for mutagenesis. The comparison of the active sites of NfsA and NfsB (Section 6.2.1), would indicate that mutations of Gly130 and Gly131 in NfsA would likely not be successful as mutations of the homologous Glu165 and Gly166 in NfsB were not (Grove *et al.*, 2003). Three areas are identified here, including ten residues to target for mutational studies (Figure 6-4). These fall into three areas. Firstly, the series of serines from 39-41, which include residues homologous to Ser40 and Thr41 in NfsB that have been mutated in improved mutants of NfsB (Grove *et al.*, 2003). Secondly, the residues interacting with the diketone ring of the FMN (Arg15, Gln 67 and His69), which include a residue (Gln67) homologous to the Asn71 that is mutated in the NfsB mutants presented in Chapter 3, which has been shown to alter the redox potential of the FMN (unpublished work by Dr Ross Anderson, University of Bristol, UK). Finally, the key Arg 225 (Section 5.13.7) and surrounding residues (Glu99, Arg133 and Aln134), as the effect of mutating Glu99 has already been shown to change substrate specificity in NfsA (Section 5.13.3 (Zenno *et al.*, 1998)). Arg225 may prove to be too important to mutate and is one of three of the ten residues suggested to bind NADPH (along with Ser41 and Gln67, which have successful homologous NfsB mutants). If looking at larger prodrugs then the flexible loop (200 - 210) would also be a potential area to target.

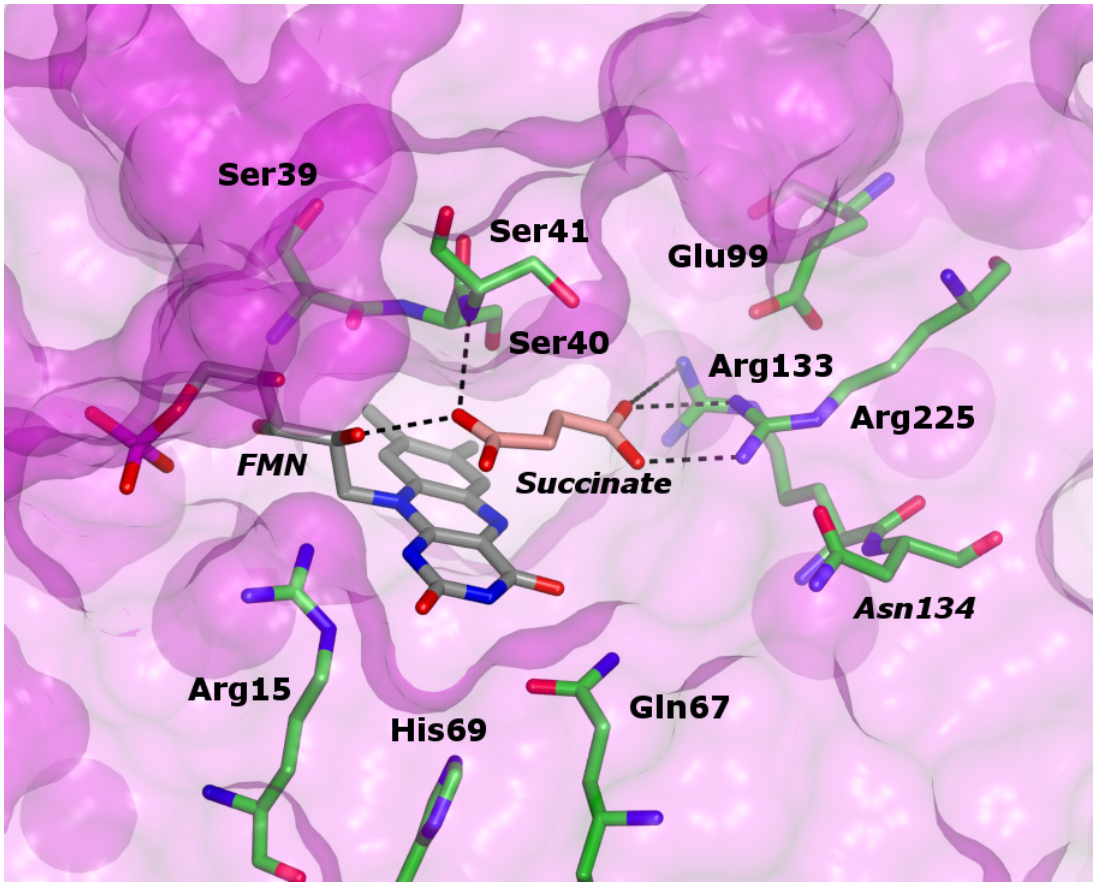


Figure 6-4: Potential residues in NfsA to be targeted for mutagenesis aiming to improve specificity for CB1954. The NfsA Succinate 1 structure is shown, with succinate in pink, FMN in grey and the residues in green. Figure produced using CCP4 Molecular Graphics.

Appendix: Solving a Crystal Structure

In protein X-ray crystallography a protein crystal is placed in an X-ray beam, and the diffracted X-rays are analysed on an area detector. The X-rays diffract of the electron clouds of the protein atoms within the crystal. This is possible as X-rays have a wavelength of the same order of magnitude of the gaps between atoms (~0.5 – 1.5 Å). The diffraction patterns formed are made up of reflections of differing intensities. When a series of diffraction patterns are collected whilst rotating the crystal, the data is processed, and each of the reflections is indexed and integrated to give a list of reflections and their intensities. Indexing is the process by which the space group is identified. The space group defines the arrangement of protein molecules in the unit cell. Symmetry operations describe the relationships between the molecules, in protein molecules only rotation-axis and screw-axis are possible, and as mirror symmetry or inversion centres are not possible due to the chirality of amino acids found in the body. Of the two possible stereoisomers only L-amino acids are found in nature. The intensities are merged to create a single file, and scaled so that reflections from different images are given accurate intensities.

In order to convert the scaled reflections into an electron density map we need to calculate the structure factor, $F_{(h,k,l)}$ for the structure. The structure factor is defined by the equation below:

$$F_{(h,k,l)} = \sum_{j=1}^{\text{atoms}} f_{(j)} \exp\left[2\pi \cdot i(hx_{(j)} + ky_{(j)} + lz_{(j)}\right]$$

The structure factor is a complex number comprising a real and imaginary component (expressed in terms of *i*), where $f_{(j)}$ is the scattering factor of atom *j*, which depends on the atom type and the diffraction angle of the reflection (*h,k,l)*. The *x, y* and *z* represent the fractional coordinates of each atom in the suspension, and *h, k,* and *l* the indices of the corresponding reflection. What we have from the scaled reflections and intensity for each reflection, and this is directly proportional to the square of the structure factor amplitude, $|F_{(h,k,l)}|$, as seen below:

$$I_{hkl} = \left|F_{(h,k,l)}\right|^2 \cdot LP \cdot A$$

Where *LP* is a combined geology and polarisation factor, dependent upon the experimental setup, and A is an absorption correction factor. From this we can see that the structure factor cannot be directly calculated from the intensities, only the amplitude of it can. This gives rise to the phase problem. The structure factors are calculated by performing a Fourier Transform, which gives the equation below comprising a magnitude ($|F_{(h,k,l)}|$) and a phase ($\varphi_{(h,k,l)}$):

$$F_{(h,k,l)} = \left|F_{(h,k,l)}\right| e^{i\varphi_{(h,k,l)}}$$

There are several ways of solving the phase problem. Molecular replacement was used in this study, and this involves using a search model of a related structure as the basis

to find the phase. Multiple/Single Wavelength Anomalous Diffraction (MAD/SAD) involves collecting data at wavelengths around the absorption edge of an atom, and phasing off the atoms signal to discover their location and give the initial phases. Finally heavy atoms can be introduced to allow a similar experiment to be performed, using the differences in the observed scattering factors to discover the position of the heavy atoms.

Once we have the structure factor the electron density can be determined through the following equation:

$$\rho_{(x,y,z)} = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{(h,k,l)}\exp[-2\pi \cdot i(hx + ky + lz)]$$

With the electron density a model structure can be determined. The fit of the model to the density is measured by the crystallographic *R*-factor:

$$R = \frac{\sum ||F_{obs}| - |F_{cal}||}{\sum |F_{obs}|}$$

Where F_{obs} is the observed structure factor, and F_{cal} is the calculated or expected structure factor. Another *R*-factor, called R_{free}, is calculated from 5% of the data left out of the structure calculations to provide an unbiased estimate of the fit of the model to the data.

Model building is performed by comparing the model to the electron density. The classic electron density map is a $2F_o$-F_c map, whilst a difference map can be produced

which highlights the differences between the model and the density producing positive (unmodelled) density and negative density (no density to support the model). This difference map is a F_o-F_c map.

Some parts of the structure are more flexible than others. The degree of movement of each atom can be judged by looking at the *B*-factor (also known as the Temperature factor or Debye–Waller factor):

$$B = 8\pi^2\langle u^2\rangle$$

Where $\langle u^2\rangle$ is the *mean squared displacement*. Each atom is given a value with units in $Å^2$, with higher values indicating more mobile parts of the structure and lower values more well-ordered parts of the structure.

References

ABUKHADER, M. M., HEAP, J., DE MATTEIS, C. I., DOUGHTY, S. W., MINTON, N. & PAOLI, M. (2007) Crystallization and preliminary X-ray characterization of the Bacillus amyloliquefaciens YwrO enzyme. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 63**,** 746-50.

ADAMS, M. J., ELLIS, G. H., GOVER, S., NAYLOR, C. E. & PHILLIPS, C. (1994) Crystallographic study of coenzyme, coenzyme analogue and substrate binding in 6-phosphogluconate dehydrogenase: implications for NADP specificity and the enzyme mechanism. *Structure,* 2**,** 651-68.

ADAMS, P. D., AFONINE, P. V., BUNKOCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

ADAMS, P. D., GROSSE-KUNSTLEVE, R. W., HUNG, L. W., IOERGER, T. R., MCCOY, A. J., MORIARTY, N. W., READ, R. J., SACCHETTINI, J. C., SAUTER, N. K. & TERWILLIGER, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr,* 58**,** 1948-54.

ANLEZARK, G. M., MELTON, R. G., SHERWOOD, R. F., COLES, B., FRIEDLOS, F. & KNOX, R. J. (1992) The bioactivation of 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB1954)--I. Purification and properties of a nitroreductase enzyme from Escherichia coli--a potential enzyme for antibody-directed enzyme prodrug therapy (ADEPT). *Biochem Pharmacol,* 44**,** 2289-95.

ANLEZARK, G. M., MELTON, R. G., SHERWOOD, R. F., WILSON, W. R., DENNY, W. A., PALMER, B. D., KNOX, R. J., FRIEDLOS, F. & WILLIAMS, A. (1995) Bioactivation of dinitrobenzamide mustards by an E. coli B nitroreductase. *Biochem Pharmacol,* 50**,** 609-18.

ASNIS, R. E., GLICK, M. C. & FRITZ, M. (1957) Effect of furacin on the dissimilation of pyruvate and formate by cell-free extracts of bacteria. *J Biol Chem,* 227**,** 863-9.

ATWELL, G. J., BOYD, M., PALMER, B. D., ANDERSON, R. F., PULLEN, S. M., WILSON, W. R. & DENNY, W. A. (1996) Synthesis and evaluation of 4-substituted analogues of 5-[N,N-bis (2-chloroethyl)amino]-2-nitrobenzamide as bioreductively activated prodrugs using an Escherichia coli nitroreductase. *Anticancer Drug Des,* 11**,** 553-67.

ATWELL, G. J., YANG, S., PRUIJN, F. B., PULLEN, S. M., HOGG, A., PATTERSON, A. V., WILSON, W. R. & DENNY, W. A. (2007) Synthesis and structure-activity relationships for 2,4-dinitrobenzamide-5-mustards as prodrugs for the Escherichia coli nfsB nitroreductase in gene therapy. *J Med Chem,* 50**,** 1197-212.

AUSTIN, E. A. & HUBER, B. E. (1993) A first step in the development of gene therapy for colorectal carcinoma: cloning, sequencing, and expression of Escherichia coli cytosine deaminase. *Mol Pharmacol,* 43**,** 380-7.

BAGSHAWE, K. D. (1987) Antibody directed enzymes revive anti-cancer prodrugs concept. *Br J Cancer,* 56**,** 531-2.

BAGSHAWE, K. D. (1990) Antibody-directed enzyme/prodrug therapy (ADEPT). *Biochem Soc Trans,* 18**,** 750-2.

BAGSHAWE, K. D., SPRINGER, C. J., SEARLE, F., ANTONIW, P., SHARMA, S. K., MELTON, R. G. & SHERWOOD, R. F. (1988) A cytotoxic agent can be generated selectively at cancer sites. *Br J Cancer,* 58**,** 700-3.

BAILEY, S. M., KNOX, R. J., HOBBS, S. M., JENKINS, T. C., MAUGER, A. B., MELTON, R. G., BURKE, P. J., CONNORS, T. A. & HART, I. R. (1996) Investigation of alternative prodrugs for use with E. coli nitroreductase in 'suicide gene' approaches to cancer therapy. *Gene Ther,* 3**,** 1143-50.

BARAK, Y., THORNE, S. H., ACKERLEY, D. F., LYNCH, S. V., CONTAG, C. H. & MATIN, A. (2006) New enzyme for reductive cancer chemotherapy, YieF, and its improvement by directed evolution. *Mol Cancer Ther,* 5**,** 97-103.

BARBOSA, T. M. & LEVY, S. B. (2002) Activation of the Escherichia coli nfnB gene by MarA through a highly divergent marbox in a class II promoter. *Mol Microbiol,* 45**,** 191-202.

BARNA, T. M., KHAN, H., BRUCE, N. C., BARSUKOV, I., SCRUTTON, N. S. & MOODY, P. C. (2001) Crystal structure of pentaerythritol tetranitrate reductase: "flipped" binding geometries for steroid substrates in different redox states of the enzyme. *J Mol Biol,* 310**,** 433-47.

BATTYE, T. G., KONTOGIANNIS, L., JOHNSON, O., POWELL, H. R. & LESLIE, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr,* 67**,** 271-81.

BELENKY, P., BOGAN, K. L. & BRENNER, C. (2007) NAD+ metabolism in health and disease. *Trends Biochem Sci,* 32**,** 12-9.

BENOUCHAN, M., DO NASCIMENTO, F., PERRET, G. Y. & COLOMBO, B. M. (2006) Delivery of the bacterial nitroreductase gene into endothelial cells prolongs the survival of tumour-bearing mice by bystander mechanisms. *Int J Oncol,* 28**,** 457-62.

BENOUCHAN, M., DO NASCIMENTO, F., SEBBAH-LOURIKI, M., SALZMANN, J. L., CREPIN, M., PERRET, G. Y. & COLOMBO, B. M. (2003) Bystander cell killing spreading from endothelial to tumor cells in a three-dimensional multicellular nodule model after Escherichia coli nitroreductase gene delivery. *Biochem Biophys Res Commun,* 311**,** 822-8.

BERNE, C., BETANCOR, L., LUCKARIFT, H. R. & SPAIN, J. C. (2006) Application of a microfluidic reactor for screening cancer prodrug activation using silica-immobilized nitrobenzene nitroreductase. *Biomacromolecules,* 7**,** 2631-6.

BIGNAMI, G. S., SENTER, P. D., GROTHAUS, P. G., FISCHER, K. J., HUMPHREYS, T. & WALLACE, P. M. (1992) N-(4'-hydroxyphenylacetyl)palytoxin: a palytoxin prodrug that can be activated by a monoclonal antibody-penicillin G amidase conjugate. *Cancer Res,* 52**,** 5759-64.

BJORNBERG, O., ROWLAND, P., LARSEN, S. & JENSEN, K. F. (1997) Active site of dihydroorotate dehydrogenase A from Lactococcus lactis investigated by chemical modification and mutagenesis. *Biochemistry,* 36**,** 16197-205.

BLAKE, C. C., KOENIG, D. F., MAIR, G. A., NORTH, A. C., PHILLIPS, D. C. & SARMA, V. R. (1965) Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature,* 206**,** 757-61.

BLATTNER, F. R., PLUNKETT, G., 3RD, BLOCH, C. A., PERNA, N. T., BURLAND, V., RILEY, M., COLLADO-VIDES, J., GLASNER, J. D., RODE, C. K., MAYHEW, G. F., GREGOR, J., DAVIS, N. W., KIRKPATRICK, H. A., GOEDEN, M. A., ROSE,

D. J., MAU, B. & SHAO, Y. (1997) The complete genome sequence of Escherichia coli K-12. *Science,* 277**,** 1453-62.

BOLAND, M. P., KNOX, R. J. & ROBERTS, J. J. (1991) The differences in kinetics of rat and human DT diaphorase result in a differential sensitivity of derived cell lines to CB 1954 (5-(aziridin-1-yl)-2,4-dinitrobenzamide). *Biochem Pharmacol,* 41**,** 867-75.

BOLIN, J. T., FILMAN, D. J., MATTHEWS, D. A., HAMLIN, R. C. & KRAUT, J. (1982) Crystal structures of Escherichia coli and Lactobacillus casei dihydrofolate reductase refined at 1.7 A resolution. I. General features and binding of methotrexate. *J Biol Chem,* 257**,** 13650-62.

BRADFORD, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem,* 72**,** 248-54.

BRIDGEWATER, J. A., KNOX, R. J., PITTS, J. D., COLLINS, M. K. & SPRINGER, C. J. (1997) The bystander effect of the nitroreductase/CB1954 enzyme/prodrug system is due to a cell-permeable metabolite. *Hum Gene Ther,* 8**,** 709-17.

BRIDGEWATER, J. A., SPRINGER, C. J., KNOX, R. J., MINTON, N. P., MICHAEL, N. P. & COLLINS, M. K. (1995) Expression of the bacterial nitroreductase enzyme in mammalian cells renders them selectively sensitive to killing by the prodrug CB1954. *Eur J Cancer,* 31A**,** 2362-70.

BROWN, N. L. & LEMOINE, N. R. (2004) Clinical trials with GDEPT: cytosine deaminase and 5-fluorocytosine. *Methods Mol Med,* 90**,** 451-7.

BRYANT, C., HUBBARD, L. & MCELROY, W. D. (1991) Cloning, nucleotide sequence, and expression of the nitroreductase gene from Enterobacter cloacae. *J Biol Chem,* 266**,** 4126-30.

BRYANT, D. W., MCCALLA, D. R., LEEKSMA, M. & LANEUVILLE, P. (1981) Type I nitroreductases of Escherichia coli. *Can J Microbiol,* 27**,** 81-6.

BRYSON, S., JULIEN, J. P., HYNES, R. C. & PAI, E. F. (2009) Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. *J Virol,* 83**,** 11862-75.

BUCHNER, E. F. (1897) Observations on 'the Principle of Identity.'. *Science,* 6**,** 809-10.

CHEN, L., WAXMAN, D. J., CHEN, D. & KUFE, D. W. (1996) Sensitization of human breast cancer cells to cyclophosphamide and ifosfamide by transfer of a liver cytochrome P450 gene. *Cancer Res,* 56**,** 1331-40.

CHEN, S., KNOX, R., WU, K., DENG, P. S., ZHOU, D., BIANCHET, M. A. & AMZEL, L. M. (1997) Molecular basis of the catalytic differences among DT-diaphorase of human, rat, and mouse. *J Biol Chem,* 272**,** 1437-9.

CHEN, V. B., ARENDALL, W. B., 3RD, HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr,* 66**,** 12-21.

CHOI, J. W., LEE, J., KOSUKE, N., JUNG, C. H. & KIM, J. S. (2007) Crystallization and preliminary X-ray diffraction analysis of ydjA, a minimal nitroreductase from Escherichia coli K12. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 63**,** 1064-6.

CHOI, J. W., LEE, J., NISHI, K., KIM, Y. S., JUNG, C. H. & KIM, J. S. (2008) Crystal structure of a minimal nitroreductase, ydjA, from Escherichia coli K12 with and without FMN cofactor. *J Mol Biol,* 377**,** 258-67.

CHRISTOFFERSON, A. & WILKIE, J. (2009) Mechanism of CB1954 reduction by Escherichia coli nitroreductase. *Biochem Soc Trans,* 37**,** 413-8.

CHUNG-FAYE, G., PALMER, D., ANDERSON, D., CLARK, J., DOWNES, M., BADDELEY, J., HUSSAIN, S., MURRAY, P. I., SEARLE, P., SEYMOUR, L., HARRIS, P. A., FERRY, D. & KERR, D. J. (2001) Virus-directed, enzyme prodrug therapy with nitroimidazole reductase: a phase I and pharmacokinetic study of its prodrug, CB1954. *Clin Cancer Res,* 7**,** 2662-8.

COBB, L. M., CONNORS, T. A., ELSON, L. A., KHAN, A. H., MITCHLEY, B. C., ROSS, W. C. & WHISSON, M. E. (1969) 2,4-dinitro-5-ethyleneiminobenzamide (CB 1954): a potent and selective inhibitor of the growth of the Walker carcinoma 256. *Biochem Pharmacol,* 18**,** 1519-27.

COLLABORATIVE COMPUTATIONAL PROJECT, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr,* 50**,** 760-3.

CONNOLLY, M. L. (1983) Solvent-accessible surfaces of proteins and nucleic acids. *Science,* 221**,** 709-13.

CORNISH-BOWDEN, A. (2004) *Fundamentals of enzyme kinetics, 3rd edition*, London Portland.

CORTES, M. L., DE FELIPE, P., MARTIN, V., HUGHES, M. A. & IZQUIERDO, M. (1998) Successful use of a plant gene in the treatment of cancer in vivo. *Gene Ther,* 5**,** 1499-507.

CUI, W., GUSTERSON, B. & CLARK, A. J. (1999) Nitroreductase-mediated cell ablation is very rapid and mediated by a p53-independent apoptotic pathway. *Gene Ther,* 6**,** 764-70.

DACHS, G. U., HUNT, M. A., SYDDALL, S., SINGLETON, D. C. & PATTERSON, A. V. (2009) Bystander or no bystander for gene directed enzyme prodrug therapy. *Molecules,* 14**,** 4517-45.

DENNY, W. A. (2002) Nitroreductase-based GDEPT. *Curr Pharm Des,* 8**,** 1349-61.

DJEHA, A. H., HULME, A., DEXTER, M. T., MOUNTAIN, A., YOUNG, L. S., SEARLE, P. F., KERR, D. J. & WRIGHTON, C. J. (2000) Expression of Escherichia coli B nitroreductase in established human tumor xenografts in mice results in potent antitumoral and bystander effects upon systemic administration of the prodrug CB1954. *Cancer Gene Ther,* 7**,** 721-31.

DJEHA, A. H., THOMSON, T. A., LEUNG, H., SEARLE, P. F., YOUNG, L. S., KERR, D. J., HARRIS, P. A., MOUNTAIN, A. & WRIGHTON, C. J. (2001) Combined adenovirus-mediated nitroreductase gene delivery and CB1954 treatment: a well-tolerated therapy for established solid tumors. *Mol Ther,* 3**,** 233-40.

DRABEK, D., GUY, J., CRAIG, R. & GROSVELD, F. (1997) The expression of bacterial nitroreductase in transgenic mice results in specific cell killing by the prodrug CB1954. *Gene Ther,* 4**,** 93-100.

EMPTAGE, C. D., KNOX, R. J., DANSON, M. J. & HOUGH, D. W. (2009) Nitroreductase from Bacillus licheniformis: a stable enzyme for prodrug activation. *Biochem Pharmacol,* 77**,** 21-9.

EMSLEY, P. & COWTAN, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

ENGEL, P. C. (1981) *Enzyme kinetics: The steady state approach, 2nd Edition*, Chapman and Hall.

ESCOLÁSTICO, C., MARÍA, M. D. S., CLARAMUNT, R. M., JIMENO, M. L., ALKORTA, I. & FOCES-FOCES, C. (1994) Imidazole and Benzimiadazole Addition to Quinones. Formation of *meso* and *d,l* Isomers and Crystal Structures of the *d,l*

Isomer of 2,3-Bis(benzimidazol-1'-yl)-1,4-Dihydroxybenzene. *Tetrahedron,* 50**,** 12489-12510.

ESTEVE-NUNEZ, A., CABALLERO, A. & RAMOS, J. L. (2001) Biological degradation of 2,4,6-trinitrotoluene. *Microbiol Mol Biol Rev,* 65**,** 335-52, table of contents.

FILLAT, C., CARRIO, M., CASCANTE, A. & SANGRO, B. (2003) Suicide gene therapy mediated by the Herpes Simplex virus thymidine kinase gene/Ganciclovir system: fifteen years of application. *Curr Gene Ther,* 3**,** 13-26.

FISCHER, E. (1894) Synthesen in der Zuckergruppe II. *Ber. Dtsch. Chem. Ges.,* 27**,** 3189.

FLEISCHMANN, R. D., ADAMS, M. D., WHITE, O., CLAYTON, R. A., KIRKNESS, E. F., KERLAVAGE, A. R., BULT, C. J., TOMB, J. F., DOUGHERTY, B. A., MERRICK, J. M. & ET AL. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science,* 269**,** 496-512.

FLEISSNER, M. R., BRUSTAD, E. M., KALAI, T., ALTENBACH, C., CASCIO, D., PETERS, F. B., HIDEG, K., PEUKER, S., SCHULTZ, P. G. & HUBBELL, W. L. (2009) Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc Natl Acad Sci U S A,* 106**,** 21637-42.

FRAAIJE, M. W. & MATTEVI, A. (2000) Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem Sci,* 25**,** 126-32.

FRAAIJE, M. W., VAN DEN HEUVEL, R. H., VAN BERKEL, W. J. & MATTEVI, A. (1999) Covalent flavinylation is essential for efficient redox catalysis in vanillyl-alcohol oxidase. *J Biol Chem,* 274**,** 35514-20.

FREEMAN, S. M., ABBOUD, C. N., WHARTENBY, K. A., PACKMAN, C. H., KOEPLIN, D. S., MOOLTEN, F. L. & ABRAHAM, G. N. (1993) The "bystander effect": tumor regression when a fraction of the tumor mass is genetically modified. *Cancer Res,* 53**,** 5274-83.

FRIEDLOS, F., COURT, S., FORD, M., DENNY, W. A. & SPRINGER, C. (1998) Gene-directed enzyme prodrug therapy: quantitative bystander cytotoxicity and DNA damage induced by CB1954 in cells expressing bacterial nitroreductase. *Gene Ther,* 5**,** 105-12.

FRIEDLOS, F., DENNY, W. A., PALMER, B. D. & SPRINGER, C. J. (1997) Mustard prodrugs for activation by Escherichia coli nitroreductase in gene-directed enzyme prodrug therapy. *J Med Chem,* 40**,** 1270-5.

FRIEDLOS, F., JARMAN, M., DAVIES, L. C., BOLAND, M. P. & KNOX, R. J. (1992) Identification of novel reduced pyridinium derivatives as synthetic co-factors for the enzyme DT diaphorase (NAD(P)H dehydrogenase (quinone), EC 1.6.99.2). *Biochem Pharmacol,* 44**,** 25-31.

GALANTE, Y. M. & HATEFI, Y. (1978) Resolution of complex I and isolation of NADH dehydrogenase and an iron--sulfur protein. *Methods Enzymol,* 53**,** 15-21.

GHISLA, S. & MASSEY, V. (1989) Mechanisms of flavoprotein-catalyzed reactions. *Eur J Biochem,* 181**,** 1-17.

GOUET, P., COURCELLE, E., STUART, D. I. & METOZ, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics,* 15**,** 305-8.

GREEN, N. K., YOUNGS, D. J., NEOPTOLEMOS, J. P., FRIEDLOS, F., KNOX, R. J., SPRINGER, C. J., ANLEZARK, G. M., MICHAEL, N. P., MELTON, R. G., FORD, M. J., YOUNG, L. S., KERR, D. J. & SEARLE, P. F. (1997) Sensitization of colorectal and pancreatic cancer cell lines to the prodrug 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB1954) by retroviral transduction and expression of the E. coli nitroreductase gene. *Cancer Gene Ther,* 4**,** 229-38.

GROHMANN, M., PAULMANN, N., FLEISCHHAUER, S., VOWINCKEL, J., PRILLER, J. & WALTHER, D. J. (2009) A mammalianized synthetic nitroreductase gene for high-level expression. *BMC Cancer,* 9**,** 301.

GROVE, J. I., LOVERING, A. L., GUISE, C., RACE, P. R., WRIGHTON, C. J., WHITE, S. A., HYDE, E. I. & SEARLE, P. F. (2003) Generation of Escherichia coli nitroreductase mutants conferring improved cell sensitization to the prodrug CB1954. *Cancer Res,* 63**,** 5532-7.

GUISE, C. P., GROVE, J. I., HYDE, E. I. & SEARLE, P. F. (2007) Direct positive selection for improved nitroreductase variants using SOS triggering of bacteriophage lambda lytic cycle. *Gene Ther,* 14**,** 690-8.

HAISMA, H. J., BOVEN, E., VAN MUIJEN, M., DE VRIES, R. & PINEDO, H. M. (1992) Analysis of a conjugate between anti-carcinoembryonic antigen monoclonal antibody and alkaline phosphatase for specific activation of the prodrug etoposide phosphate. *Cancer Immunol Immunother,* 34**,** 343-8.

HAJOS, A. K. & WINSTON, G. W. (1991) Purified NAD(P)H-quinone oxidoreductase enhances the mutagenicity of dinitropyrenes in vitro. *J Biochem Toxicol,* 6**,** 277-82.

HAMMILL, J. T., MIYAKE-STONER, S., HAZEN, J. L., JACKSON, J. C. & MEHL, R. A. (2007) Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat Protoc,* 2**,** 2601-7.

HAPKE, D. M., STEGMANN, A. P. & MITCHELL, B. S. (1996) Retroviral transfer of deoxycytidine kinase into tumor cell lines enhances nucleoside toxicity. *Cancer Res,* 56**,** 2343-7.

HAY, M. P., SYKES, B. M., DENNY, W. A. & WILSON, W. R. (1999a) A 2-nitroimidazole carbamate prodrug of 5-amimo-1-(chloromethyl)-3-[(5,6,7-trimethoxyindol-2-yl)carbony l]-1,2-dihydro-3H--benz[E]indole (amino-seco-CBI-TMI) for use with ADEPT and GDEPT. *Bioorg Med Chem Lett,* 9**,** 2237-42.

HAY, M. P., WILSON, W. R. & DENNY, W. A. (1995) A Novel Enediyne Prodrug for Antibody-Directed Enzyme Prodrug Therapy (ADEPT) using *E. coli* B Nitroreductase. *Bioorg. Med. Chem. Lett.,* 5**,** 2829-2834.

HAY, M. P., WILSON, W. R. & DENNY, W. A. (1999b) Nitrobenzyl carbamate prodrugs of enediynes for nitroreductase gene-directed enzyme prodrug therapy (GDEPT). *Bioorg Med Chem Lett,* 9**,** 3417-22.

HAYNES, C. A., KODER, R. L., MILLER, A. F. & RODGERS, D. W. (2002) Structures of nitroreductase in three states: effects of inhibitor binding and reduction. *J Biol Chem,* 277**,** 11513-20.

HELSBY, N. A., FERRY, D. M., PATTERSON, A. V., PULLEN, S. M. & WILSON, W. R. (2004) 2-Amino metabolites are key mediators of CB 1954 and SN 23862 bystander effects in nitroreductase GDEPT. *Br J Cancer,* 90**,** 1084-92.

HELSBY, N. A., WHEELER, S. J., PRUIJN, F. B., PALMER, B. D., YANG, S., DENNY, W. A. & WILSON, W. R. (2003) Effect of nitroreduction on the alkylating reactivity and cytotoxicity of the 2,4-dinitrobenzamide-5-aziridine CB 1954 and the corresponding nitrogen mustard SN 23862: distinct mechanisms of bioreductive activation. *Chem Res Toxicol,* 16**,** 469-78.

HENDRIX, R. W., ROBERTS, J. W., STAHL, F. W. & WEISBERG, R. A. (Eds.) (1983) *Lambda II,* New York, Cold Spring Harbor Laboratory Press.

HIRS, C. H., MOORE, S. & STEIN, W. H. (1960) The sequence of the amino acid residues in performic acid-oxidized ribonuclease. *J Biol Chem,* 235**,** 633-47.

HU, L., WU, X., HAN, J., CHEN, L., VASS, S. O., BROWNE, P., HALL, B. S., BOT, C., GOBALAKRISHNAPILLAI, V., SEARLE, P. F., KNOX, R. J. & WILKINSON, S. R. (2011) Synthesis and structure-activity relationships of nitrobenzyl phosphoramide mustards as nitroreductase-activated prodrugs. *Bioorg Med Chem Lett,* 21**,** 3986-91.

HUANG, C. H., LAI, W. L., LEE, M. H., CHEN, C. J., VASELLA, A., TSAI, Y. C. & LIAW, S. H. (2005) Crystal structure of glucooligosaccharide oxidase from Acremonium strictum: a novel flavinylation of 6-S-cysteinyl, 8alpha-N1-histidyl FAD. *J Biol Chem,* 280**,** 38831-8.

HUBER, B. E., RICHARDS, C. A. & AUSTIN, E. A. (1994) Virus-directed enzyme/prodrug therapy (VDEPT). Selectively engineering drug sensitivity into tumors. *Ann N Y Acad Sci,* 716**,** 104-14; discussion 140-3.

JABERIPOUR, M., VASS, S. O., GUISE, C. P., GROVE, J. I., KNOX, R. J., HU, L., HYDE, E. I. & SEARLE, P. F. (2010) Testing double mutants of the enzyme nitroreductase for enhanced cell sensitisation to prodrugs: effects of combining beneficial single mutations. *Biochem Pharmacol,* 79**,** 102-11.

JACKSON, J. C., DUFFY, S. P., HESS, K. R. & MEHL, R. A. (2006) Improving nature's enzyme active site with genetically encoded unnatural amino acids. *J Am Chem Soc,* 128**,** 11124-7.

JARROM, D. (2008) Optimisation of the *E. coli* nitroreductases for gene-directed enzyme prodrug therapy. *School of Biosciences.* Birmingham, UK, University of Birmingham.

JARROM, D., JABERIPOUR, M., GUISE, C. P., DAFF, S., WHITE, S. A., SEARLE, P. F. & HYDE, E. I. (2009) Steady-state and stopped-flow kinetic studies of three Escherichia coli NfsB mutants with enhanced activity for the prodrug CB1954. *Biochemistry,* 48**,** 7665-72.

JIANG, Y., HAN, J., YU, C., VASS, S. O., SEARLE, P. F., BROWNE, P., KNOX, R. J. & HU, L. (2006) Design, synthesis, and biological evaluation of cyclic and acyclic nitrobenzylphosphoramide mustards for E. coli nitroreductase activation. *J Med Chem,* 49**,** 4333-43.

JOHANSSON, E., PARKINSON, G. N., DENNY, W. A. & NEIDLE, S. (2003) Studies on the nitroreductase prodrug-activating system. Crystal structures of complexes with the inhibitor dicoumarol and dinitrobenzamide prodrugs and of the enzyme active form. *J Med Chem,* 46**,** 4009-20.

JOOSTEN, V. & VAN BERKEL, W. J. (2007) Flavoenzymes. *Curr Opin Chem Biol,* 11**,** 195-202.

KABSCH, W. Xds. *Acta Crystallogr D Biol Crystallogr,* 66**,** 125-32.

KERR, D. J., YOUNG, L. S., SEARLE, P. F. & MCNEISH, I. A. (1997) Gene directed enzyme prodrug therapy for cancer. *Adv Drug Deliv Rev,* 26**,** 173-184.

KHAN, H., BARNA, T., BRUCE, N. C., MUNRO, A. W., LEYS, D. & SCRUTTON, N. S. (2005) Proton transfer in the oxidative half-reaction of pentaerythritol tetranitrate reductase. Structure of the reduced enzyme-progesterone complex and the roles of residues Tyr186, His181, His184. *Febs J,* 272**,** 4660-71.

KITTS, C. L., GREEN, C. E., OTLEY, R. A., ALVAREZ, M. A. & UNKEFER, P. J. (2000) Type I nitroreductases in soil enterobacteria reduce TNT (2,4,6,-trinitrotoluene) and RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine). *Can J Microbiol,* 46**,** 278-82.

KNOX, R. J., BOLAND, M. P., FRIEDLOS, F., COLES, B., SOUTHAN, C. & ROBERTS, J. J. (1988a) The nitroreductase enzyme in Walker cells that activates 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB 1954) to 5-(aziridin-1-yl)-4-hydroxylamino-2-

nitrobenzamide is a form of NAD(P)H dehydrogenase (quinone) (EC 1.6.99.2). *Biochem Pharmacol,* 37**,** 4671-7.

KNOX, R. J., BURKE, P. J., CHEN, S. & KERR, D. J. (2003) CB 1954: from the Walker tumor to NQO2 and VDEPT. *Curr Pharm Des,* 9**,** 2091-104.

KNOX, R. J., FRIEDLOS, F., BIGGS, P. J., FLITTER, W. D., GASKELL, M., GODDARD, P., DAVIES, L. & JARMAN, M. (1993a) Identification, synthesis and properties of 5-(aziridin-1-yl)-2-nitro-4-nitrosobenzamide, a novel DNA crosslinking agent derived from CB1954. *Biochem Pharmacol,* 46**,** 797-803.

KNOX, R. J., FRIEDLOS, F. & BOLAND, M. P. (1993b) The bioactivation of CB 1954 and its use as a prodrug in antibody-directed enzyme prodrug therapy (ADEPT). *Cancer Metastasis Rev,* 12**,** 195-212.

KNOX, R. J., FRIEDLOS, F., JARMAN, M., DAVIES, L. C., GODDARD, P., ANLEZARK, G. M., MELTON, R. G. & SHERWOOD, R. F. (1995) Virtual cofactors for an Escherichia coli nitroreductase enzyme: relevance to reductively activated prodrugs in antibody directed enzyme prodrug therapy (ADEPT). *Biochem Pharmacol,* 49**,** 1641-7.

KNOX, R. J., FRIEDLOS, F., JARMAN, M. & ROBERTS, J. J. (1988b) A new cytotoxic, DNA interstrand crosslinking agent, 5-(aziridin-1-yl)-4-hydroxylamino-2-nitrobenzamide, is formed from 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB 1954) by a nitroreductase enzyme in Walker carcinoma cells. *Biochem Pharmacol,* 37**,** 4661-9.

KNOX, R. J., FRIEDLOS, F., MARCHBANK, T. & ROBERTS, J. J. (1991) Bioactivation of CB 1954: reaction of the active 4-hydroxylamino derivative with thioesters to form the ultimate DNA-DNA interstrand crosslinking species. *Biochem Pharmacol,* 42**,** 1691-7.

KNOX, R. J., FRIEDLOS, F., SHERWOOD, R. F., MELTON, R. G. & ANLEZARK, G. M. (1992) The bioactivation of 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB1954)--II. A comparison of an Escherichia coli nitroreductase and Walker DT diaphorase. *Biochem Pharmacol,* 44**,** 2297-301.

KNOX, R. J., JENKINS, T. C., HOBBS, S. M., CHEN, S., MELTON, R. G. & BURKE, P. J. (2000) Bioactivation of 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB 1954) by human NAD(P)H quinone oxidoreductase 2: a novel co-substrate-mediated antitumor prodrug therapy. *Cancer Res,* 60**,** 4179-86.

KOBORI, T., LEE, W. C., AKAGI, T., SASAKI, H., ZENNO, S., SAIGO, K. & TANOKURA, M. (1999) Crystallization and preliminary crystallographic analysis of major nitroreductase from Escherichia coli. *Acta Crystallogr D Biol Crystallogr,* 55**,** 1901-2.

KOBORI, T., SASAKI, H., LEE, W. C., ZENNO, S., SAIGO, K., MURPHY, M. E. & TANOKURA, M. (2001) Structure and site-directed mutagenesis of a flavoprotein from Escherichia coli that reduces nitrocompounds: alteration of pyridine nucleotide binding by a single amino acid substitution. *J Biol Chem,* 276**,** 2816-23.

KODER, R. L. & MILLER, A. F. (1998) Steady-state kinetic mechanism, stereospecificity, substrate and inhibitor specificity of Enterobacter cloacae nitroreductase. *Biochim Biophys Acta,* 1387**,** 395-405.

KOIVUNEN, P., HIRSILA, M., REMES, A. M., HASSINEN, I. E., KIVIRIKKO, K. I. & MYLLYHARJU, J. (2007) Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: possible links between cell metabolism and stabilization of HIF. *J Biol Chem,* 282**,** 4524-32.

KOSHLAND, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A,* 44**,** 98-104.

KÜHNE, W. (1877) *Uber das Verhalten verschiedener organisirter und sog. ungeformter Fermente*.

LAEMMLI, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LASKOWSKI, R. A. (2009) PDBsum new things. *Nucleic Acids Res,* 37**,** D355-9.

LESLIE, A. G. W. & POWELL, H. R. (2007) Processing Diffraction Data with Mosflm. *Evolving Methods for Macromolecular Crystallography.*

LI, Z., HAN, J., JIANG, Y., BROWNE, P., KNOX, R. J. & HU, L. (2003) Nitrobenzocyclophosphamides as potential prodrugs for bioreductive activation: synthesis, stability, enzymatic reduction, and antiproliferative activity in cell culture. *Bioorg Med Chem,* 11**,** 4171-8.

LIGHTFOOT, R. T., SHUMAN, D. & ISCHIROPOULOS, H. (2000) Oxygen-insensitive nitroreductases of Escherichia coli do not reduce 3-nitrotyrosine. *Free Radic Biol Med,* 28**,** 1132-6.

LIOCHEV, S. I., HAUSLADEN, A. & FRIDOVICH, I. (1999) Nitroreductase A is regulated as a member of the soxRS regulon of Escherichia coli. *Proc Natl Acad Sci U S A,* 96**,** 3537-9.

LOVERING, A. L., HYDE, E. I., SEARLE, P. F. & WHITE, S. A. (2001) The structure of Escherichia coli nitroreductase complexed with nicotinic acid: three crystal forms at 1.7 A, 1.8 A and 2.4 A resolution. *J Mol Biol,* 309**,** 203-13.

LUDWIG, C., MICHIELS, P. J., WU, X., KAVANAGH, K. L., PILKA, E., JANSSON, A., OPPERMANN, U. & GUNTHER, U. L. (2008) SALMON: solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy. *J Med Chem,* 51**,** 1-3.

MARAIS, R., SPOONER, R. A., LIGHT, Y., MARTIN, J. & SPRINGER, C. J. (1996) Gene-directed enzyme prodrug therapy with a mustard prodrug/carboxypeptidase G2 combination. *Cancer Res,* 56**,** 4735-42.

MASON, D. E. & LIEBLER, D. C. (2000) Characterization of benzoquinone-peptide adducts by electrospray mass spectrometry. *Chem Res Toxicol,* 13**,** 976-82.

MASON, R. P. & HOLTZMAN, J. L. (1975a) The mechanism of microsomal and mitochondrial nitroreductase. Electron spin resonance evidence for nitroaromatic free radical intermediates. *Biochemistry,* 14**,** 1626-32.

MASON, R. P. & HOLTZMAN, J. L. (1975b) The role of catalytic superoxide formation in the O2 inhibition of nitroreductase. *Biochem Biophys Res Commun,* 67**,** 1267-74.

MASSEY, V. (1995) Introduction: flavoprotein structure and mechanism. *FASEB J,* 9**,** 473-5.

MAUGER, A. B., BURKE, P. J., SOMANI, H. H., FRIEDLOS, F. & KNOX, R. J. (1994) Self-immolative prodrugs: candidates for antibody-directed enzyme prodrug therapy in conjunction with a nitroreductase enzyme. *J Med Chem,* 37**,** 3452-8.

MAYHEW, S. G. (1999) The effects of pH and semiquinone formation on the oxidation-reduction potentials of flavin mononucleotide. A reappraisal. *Eur J Biochem,* 265**,** 698-702.

MCCALLA, D. R., KAISER, C. & GREEN, M. H. (1978) Genetics of nitrofurazone resistance in Escherichia coli. *J Bacteriol,* 133**,** 10-6.

MCCALLA, D. R., REUVERS, A. & KAISER, C. (1970) Mode of Action of Nitrofurazone. *J Bacteriol,* 104**,** 1126-1134.

MCCALLA, D. R., REUVERS, A. & KAISER, C. (1971) Breakage of bacterial DNA by nitrofuran derivatives. *Cancer Res,* 31**,** 2184-8.

MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr,* 40**,** 658-674.

MCNEISH, I. A., GREEN, N. K., GILLIGAN, M. G., FORD, M. J., MAUTNER, V., YOUNG, L. S., KERR, D. J. & SEARLE, P. F. (1998) Virus directed enzyme prodrug therapy for ovarian and pancreatic cancer using retrovirally delivered E. coli nitroreductase and CB1954. *Gene Ther,* 5**,** 1061-9.

MEWIES, M., PACKMAN, L. C., MATHEWS, F. S. & SCRUTTON, N. S. (1996) Flavinylation in wild-type trimethylamine dehydrogenase and differentially charged mutant enzymes: a study of the protein environment around the N1 of the flavin isoalloxazine. *Biochem J,* 317 (Pt 1)**,** 267-72.

MICHAEL, N. P., BREHM, J. K., ANLEZARK, G. M. & MINTON, N. P. (1994) Physical characterisation of the Escherichia coli B gene encoding nitroreductase and its over-expression in Escherichia coli K12. *FEMS Microbiol Lett,* 124**,** 195-202.

MIDDLETON, M. R., KNOX, R., CATTELL, E., OPPERMANN, U., MIDGLEY, R., ALI, R., AUTON, T., AGARWAL, R., ANDERSON, D., SARKER, D., JUDSON, I., OSAWA, T., SPANSWICK, V. J., DAVIES, S., HARTLEY, J. A. & KERR, D. J. (2010) Quinone oxidoreductase-2-mediated prodrug cancer therapy. *Sci Transl Med,* 2**,** 40-50.

MINTON, N. P. (2003) Clostridia in cancer therapy. *Nat Rev Microbiol,* 1**,** 237-42.

MITCHELL, D. J. & MINCHIN, R. F. (2008) E. coli nitroreductase/CB1954 gene-directed enzyme prodrug therapy: role of arylamine N-acetlytransferase 2. *Cancer Gene Ther,* 15**,** 758-64.

MONKS, T. J., HANZLIK, R. P., COHEN, G. M., ROSS, D. & GRAHAM, D. G. (1992) Quinone chemistry and toxicity. *Toxicol Appl Pharmacol,* 112**,** 2-16.

MONOD, J., CHANGEUX, J. P. & JACOB, F. (1963) Allosteric proteins and cellular control systems. *J Mol Biol,* 6**,** 306-29.

MONOD, J., WYMAN, J. & CHANGEUX, J. P. (1965) On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol,* 12**,** 88-118.

MOOLTEN, F. L. (1986) Tumor chemosensitivity conferred by inserted herpes thymidine kinase genes: paradigm for a prospective cancer control strategy. *Cancer Res,* 46**,** 5276-81.

MOSER, C. C., ANDERSON, J. L. & DUTTON, P. L. (2010) Guidelines for tunneling in enzymes. *Biochim Biophys Acta,* 1797**,** 1573-86.

MURSHUDOV, G. N., VAGIN, A. A. & DODSON, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr,* 53**,** 240-55.

NICULESCU-DUVAZ, I. & SPRINGER, C. J. (1997) Gene-directed enzyme prodrug therapy: a review of enzyme/prodrug combinations. *Expert Opin Investig Drugs,* 6**,** 685-703.

NISHIHARA, E., NAGAYAMA, Y., NARIMATSU, M., NAMBA, H., WATANABE, M., NIWA, M. & YAMASHITA, S. (1998) Treatment of thyroid carcinoma cells with four different suicide gene/prodrug combinations in vitro. *Anticancer Res,* 18**,** 1521-5.

PAGE, C. C., MOSER, C. C. & DUTTON, P. L. (2003) Mechanism for electron transfer within and between proteins. *Curr Opin Chem Biol,* 7**,** 551-6.

PALMER, B. D., WILSON, W. R., ATWELL, G. J., SCHULTZ, D., XU, X. Z. & DENNY, W. A. (1994) Hypoxia-selective antitumor agents. 9. Structure-activity relationships for hypoxia-selective cytotoxicity among analogues of 5-[N,N-bis(2-chloroethyl)amino]-2,4-dinitrobenzamide. *J Med Chem,* 37**,** 2175-84.

PALMER, B. D., WILSON, W. R., CLIFFE, S. & DENNY, W. A. (1992) Hypoxia-selective antitumor agents. 5. Synthesis of water-soluble nitroaniline mustards with selective cytotoxicity for hypoxic mammalian cells. *J Med Chem,* 35**,** 3214-22.

PALMER, B. D., WILSON, W. R., PULLEN, S. M. & DENNY, W. A. (1990) Hypoxia-selective antitumor agents. 3. Relationships between structure and cytotoxicity against cultured tumor cells for substituted N,N-bis(2-chloroethyl)anilines. *J Med Chem,* 33**,** 112-21.

PALMER, D. H., MAUTNER, V., MIRZA, D., OLIFF, S., GERRITSEN, W., VAN DER SIJP, J. R., HUBSCHER, S., REYNOLDS, G., BONNEY, S., RAJARATNAM, R., HULL, D., HORNE, M., ELLIS, J., MOUNTAIN, A., HILL, S., HARRIS, P. A., SEARLE, P. F., YOUNG, L. S., JAMES, N. D. & KERR, D. J. (2004) Virus-directed enzyme prodrug therapy: intratumoral administration of a replication-deficient adenovirus encoding nitroreductase to patients with resectable liver cancer. *J Clin Oncol,* 22**,** 1546-52.

PARKINSON, G. N., SKELLY, J. V. & NEIDLE, S. (2000) Crystal structure of FMN-dependent nitroreductase from Escherichia coli B: a prodrug-activating enzyme. *J Med Chem,* 43**,** 3624-31.

PATEL, P., YOUNG, J. G., MAUTNER, V., ASHDOWN, D., BONNEY, S., PINEDA, R. G., COLLINS, S. I., SEARLE, P. F., HULL, D., PEERS, E., CHESTER, J., WALLACE, D. M., DOHERTY, A., LEUNG, H., YOUNG, L. S. & JAMES, N. D. (2009) A phase I/II clinical trial in localized prostate cancer of an adenovirus expressing nitroreductase with CB1984. *Mol Ther,* 17**,** 1292-9.

PETERSON, F. J., MASON, R. P., HOVSEPIAN, J. & HOLTZMAN, J. L. (1979) Oxygen-sensitive and -insensitive nitroreduction by Escherichia coli and rat hepatic microsomes. *J Biol Chem,* 254**,** 4009-14.

PRATT, W. B. & RUDDON, R. W. (1979) The alkylating agents. IN PRATT, W. B. & RUDDON, R. W. (Eds.) *The Anti-cancer Drugs.* Oxford, Oxford University Press.

PROSSER, G. A., COPP, J. N., SYDDALL, S. P., WILLIAMS, E. M., SMAILL, J. B., WILSON, W. R., PATTERSON, A. V. & ACKERLEY, D. F. (2010a) Discovery and evaluation of Escherichia coli nitroreductases that activate the anti-cancer prodrug CB1954. *Biochem Pharmacol,* 79**,** 678-87.

PROSSER, G. A., PATTERSON, A. V. & ACKERLEY, D. F. (2010b) uvrB gene deletion enhances SOS chromotest sensitivity for nitroreductases that preferentially generate the 4-hydroxylamine metabolite of the anti-cancer prodrug CB1954. *J Biotechnol,* 150**,** 190-4.

RACE, P. R., LOVERING, A. L., GREEN, R. M., OSSOR, A., WHITE, S. A., SEARLE, P. F., WRIGHTON, C. J. & HYDE, E. I. (2005) Structural and mechanistic studies of Escherichia coli nitroreductase with the antibiotic nitrofurazone. Reversed binding orientations in different redox states of the enzyme. *J Biol Chem,* 280**,** 13256-64.

RACE, P. R., LOVERING, A. L., WHITE, S. A., GROVE, J. I., SEARLE, P. F., WRIGHTON, C. W. & HYDE, E. I. (2007) Kinetic and structural characterisation of

Escherichia coli nitroreductase mutants showing improved efficacy for the prodrug substrate CB1954. *J Mol Biol,* 368**,** 481-92.

RAINOV, N. G., DOBBERSTEIN, K. U., SENA-ESTEVES, M., HERRLINGER, U., KRAMM, C. M., PHILPOT, R. M., HILTON, J., CHIOCCA, E. A. & BREAKEFIELD, X. O. (1998) New prodrug activation gene therapy for cancer using cytochrome P450 4B1 and 2-aminoanthracene/4-ipomeanol. *Hum Gene Ther,* 9**,** 1261-73.

RAND, T., QVIST, K. B., WALTER, C. P. & POULSEN, C. H. (2006) Characterization of the flavin association in hexose oxidase from Chondrus crispus. *Febs J,* 273**,** 2693-703.

REISS, P. D., ZUURENDONK, P. F. & VEECH, R. L. (1984) Measurement of tissue purine, pyrimidine, and other nucleotides by radial compression high-performance liquid chromatography. *Anal Biochem,* 140**,** 162-71.

REYES, V. M., SAWAYA, M. R., BROWN, K. A. & KRAUT, J. (1995) Isomorphous crystal structures of Escherichia coli dihydrofolate reductase complexed with folate, 5-deazafolate, and 5,10-dideazatetrahydrofolate: mechanistic implications. *Biochemistry,* 34**,** 2710-23.

ROBERTS, J. J., FRIEDLOS, F. & KNOX, R. J. (1986) CB 1954 (2,4-dinitro-5-aziridinyl benzamide) becomes a DNA interstrand crosslinking agent in Walker tumour cells. *Biochem Biophys Res Commun,* 140**,** 1073-8.

ROBERTS, J. W. & DEVORET, R. (1983) Lysogenic induction. IN HENDRIX, R. W., ROBERTS, J. W., STAHL, F. W. & WEISBERG, R. A. (Eds.) *Lambda II.* New York, Cold Spring Harbor Laboratory Press.

ROFFLER, S. R., WANG, S. M., CHERN, J. W., YEH, M. Y. & TUNG, E. (1991) Anti-neoplastic glucuronide prodrug treatment of human tumor cells targeted with a monoclonal antibody-enzyme conjugate. *Biochem Pharmacol,* 42**,** 2062-5.

ROLDAN, M. D., PEREZ-REINADO, E., CASTILLO, F. & MORENO-VIVIAN, C. (2008) Reduction of polynitroaromatic compounds: the bacterial nitroreductases. *FEMS Microbiol Rev,* 32**,** 474-500.

RUSSELL, P. J. & KHATRI, A. (2006) Novel gene-directed enzyme prodrug therapies against prostate cancer. *Expert Opin Investig Drugs,* 15**,** 947-61.

SAGNOU, M. J., HOWARD, P. W., GREGSON, S. J., ENO-AMOOQUAYE, E., BURKE, P. J. & THURSTON, D. E. (2000) Design and synthesis of novel pyrrolobenzodiazepine (PBD) prodrugs for ADEPT and GDEPT. *Bioorg Med Chem Lett,* 10**,** 2083-6.

SASSANFAR, M. & ROBERTS, J. W. (1990) Nature of the SOS-inducing signal in Escherichia coli. The involvement of DNA replication. *J Mol Biol,* 212**,** 79-96.

SASTRY, S. S. & JAYARAMAN, R. (1984) Nitrofurantoin-resistant mutants of Escherichia coli: isolation and mapping. *Mol Gen Genet,* 196**,** 379-80.

SATCHI, R., CONNORS, T. A. & DUNCAN, R. (2001) PDEPT: polymer-directed enzyme prodrug therapy. I. HPMA copolymer-cathepsin B and PK1 as a model combination. *Br J Cancer,* 85**,** 1070-6.

SEARLE, P. F., CHEN, M. J., HU, L., RACE, P. R., LOVERING, A. L., GROVE, J. I., GUISE, C., JABERIPOUR, M., JAMES, N. D., MAUTNER, V., YOUNG, L. S., KERR, D. J., MOUNTAIN, A., WHITE, S. A. & HYDE, E. I. (2004) Nitroreductase: a prodrug-activating enzyme for cancer gene therapy. *Clin Exp Pharmacol Physiol,* 31**,** 811-6.

SENTER, P. D., WALLACE, P. M., SVENSSON, H. P., KERR, D. E., HELLSTROM, I. & HELLSTROM, K. E. (1991) Activation of prodrugs by antibody-enzyme conjugates. *Adv Exp Med Biol,* 303**,** 97-105.

SINGLETON, D. C., LI, D., BAI, S. Y., SYDDALL, S. P., SMAILL, J. B., SHEN, Y., DENNY, W. A., WILSON, W. R. & PATTERSON, A. V. (2007) The nitroreductase prodrug SN 28343 enhances the potency of systemically administered armed oncolytic adenovirus ONYX-411(NTR). *Cancer Gene Ther,* 14**,** 953-67.

SKELLY, J. V., KNOX, R. J. & JENKINS, T. C. (2001) Aerobic nitroreduction by flavoproteins: enzyme structure, mechanisms and role in cancer chemotherapy. *Mini Rev Med Chem,* 1**,** 293-306.

SMITH, G. N. & WORREL, C. S. (1949) Enzymatic reduction of chloramphenicol. *Arch Biochem,* 24**,** 216-23.

SPAIN, J. C. (1995) Biodegradation of nitroaromatic compounds. *Annu Rev Microbiol,* 49**,** 523-55.

STUDIER, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif,* 41**,** 207-34.

SUMNER, J. B. (1926) The Isolation and Crystallization of the Enzyme Urease. Preliminary Paper *J. Biol. Chem.,* 69**,** 435-441.

SUNTERS, A., SPRINGER, C. J., BAGSHAWE, K. D., SOUHAMI, R. L. & HARTLEY, J. A. (1992) The cytotoxicity, DNA crosslinking ability and DNA sequence selectivity of the aniline mustards melphalan, chlorambucil and 4-[bis(2-chloroethyl)amino] benzoic acid. *Biochem Pharmacol,* 44**,** 59-64.

TANG, M. H., HELSBY, N. A., WILSON, W. R. & TINGLE, M. D. (2005) Aerobic 2- and 4-nitroreduction of CB 1954 by human liver. *Toxicology,* 216**,** 129-39.

TANNER, J. J., LEI, B., TU, S. C. & KRAUSE, K. L. (1996) Flavin reductase P: structure of a dimeric enzyme that reduces flavin. *Biochemistry,* 35**,** 13531-9.

TERCEL, M., DENNY, W. A. & WILSON, W. R. (1996) A Novel Nitro-Substituted *seco*-CI: Application as a Reductively Activated ADEPT Prodrug. *Bioorg. Med. Chem. Lett.,* 6**,** 2741-2744.

THATCHER, N. J., EDWARDS, R. J., LEMOINE, N. R., DOEHMER, J. & DAVIES, D. S. (2000) The potential of acetaminophen as a prodrug in gene-directed enzyme prodrug therapy. *Cancer Gene Ther,* 7**,** 521-5.

THEOBALD, D. L. & WUTTKE, D. S. (2008) Accurate structural correlations from maximum likelihood superpositions. *PLoS Comput Biol,* 4**,** e43.

THEYS, J., PENNINGTON, O., DUBOIS, L., ANLEZARK, G., VAUGHAN, T., MENGESHA, A., LANDUYT, W., ANNE, J., BURKE, P. J., DURRE, P., WOUTERS, B. G., MINTON, N. P. & LAMBIN, P. (2006) Repeated cycles of Clostridium-directed enzyme prodrug therapy result in sustained antitumour effects in vivo. *Br J Cancer,* 95**,** 1212-9.

THORN, J. M., BARTON, J. D., DIXON, N. E., OLLIS, D. L. & EDWARDS, K. J. (1995) Crystal structure of Escherichia coli QOR quinone oxidoreductase complexed with NADPH. *J Mol Biol,* 249**,** 785-99.

TIETZE, L. F. & SCHMUCK, K. (2011) Prodrugs for targeted tumor therapies: recent developments in ADEPT, GDEPT and PMT. *Curr Pharm Des,* 17**,** 3527-47.

TONOMURA, B., NAKATANI, H., OHNISHI, M., YAMAGUCHI-ITO, J. & HIROMI, K. (1978) Test reactions for a stopped-flow apparatus. Reduction of 2,6-dichlorophenolindophenol and potassium ferricyanide by L-ascorbic acid. *Anal Biochem,* 84**,** 370-83.

TYCHOPOULOS, M., CORCOS, L., GENNE, P., BEAUNE, P. & DE WAZIERS, I. (2005) A virus-directed enzyme prodrug therapy (VDEPT) strategy for lung cancer using a CYP2B6/NADPH-cytochrome P450 reductase fusion protein. *Cancer Gene Ther,* 12**,** 497-508.

UEDA, O., KITAMURA, S., OHASHI, K., SUGIHARA, K. & OHTA, S. (2003) Xanthine oxidase-catalyzed metabolism of 2-nitrofluorene, a carcinogenic air pollutant, in rat skin. *Drug Metab Dispos,* 31**,** 367-72.

URZHUMTSEVA, L., AFONINE, P. V., ADAMS, P. D. & URZHUMTSEV, A. (2009) Crystallographic model quality at a glance. *Acta Crystallogr D Biol Crystallogr,* 65**,** 297-300.

VASS, S. O., JARROM, D., WILSON, W. R., HYDE, E. I. & SEARLE, P. F. (2009) E. coli NfsA: an alternative nitroreductase for prodrug activation gene therapy in combination with CB1954. *Br J Cancer,* 100**,** 1903-11.

VENITT, S. & CROFTON-SLEIGH, C. (1987) The toxicity and mutagenicity of the anti-tumour drug 5-aziridino-2,4-dinitrobenzamide (CB1954) is greatly reduced in a nitroreductase-deficient strain of E. coli. *Mutagenesis,* 2**,** 375-81.

WATANABE, M., ISHIDATE, M., JR. & NOHMI, T. (1990) Nucleotide sequence of Salmonella typhimurium nitroreductase gene. *Nucleic Acids Res,* 18**,** 1059.

WEEDON, S. J., GREEN, N. K., MCNEISH, I. A., GILLIGAN, M. G., MAUTNER, V., WRIGHTON, C. J., MOUNTAIN, A., YOUNG, L. S., KERR, D. J. & SEARLE, P. F. (2000) Sensitisation of human carcinoma cells to the prodrug CB1954 by adenovirus vector-mediated expression of E. coli nitroreductase. *Int J Cancer,* 86**,** 848-54.

WESTPHAL, E. M., GE, J., CATCHPOLE, J. R., FORD, M. & KENNEY, S. C. (2000) The nitroreductase/CB1954 combination in Epstein-Barr virus-positive B-cell lines: induction of bystander killing in vitro and in vivo. *Cancer Gene Ther,* 7**,** 97-106.

WHITEWAY, J., KOZIARZ, P., VEALL, J., SANDHU, N., KUMAR, P., HOECHER, B. & LAMBERT, I. B. (1998) Oxygen-insensitive nitroreductases: analysis of the roles of nfsA and nfsB in development of resistance to 5-nitrofuran derivatives in Escherichia coli. *J Bacteriol,* 180**,** 5529-39.

WILSON, D. K., BOHREN, K. M., GABBAY, K. H. & QUIOCHO, F. A. (1992) An unlikely sugar substrate site in the 1.65 A structure of the human aldose reductase holoenzyme implicated in diabetic complications. *Science,* 257**,** 81-4.

WILSON, W. R., PULLEN, S. M., HOGG, A., HELSBY, N. A., HICKS, K. O. & DENNY, W. A. (2002) Quantitation of bystander effects in nitroreductase suicide gene therapy using three-dimensional cell cultures. *Cancer Res,* 62**,** 1425-32.

WINKLER, A., HARTNER, F., KUTCHAN, T. M., GLIEDER, A. & MACHEROUX, P. (2006) Biochemical evidence that berberine bridge enzyme belongs to a novel family of flavoproteins containing a bi-covalently attached FAD cofactor. *J Biol Chem,* 281**,** 21276-85.

WORD, J. M., LOVELL, S. C., RICHARDSON, J. S. & RICHARDSON, D. C. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol,* 285**,** 1735-47.

WORKMAN, P., MORGAN, J. E., TALBOT, K., WRIGHT, K. A., DONALDSON, J. & TWENTYMAN, P. R. (1986) CB 1954 revisited. II. Toxicity and antitumour activity. *Cancer Chemother Pharmacol,* 16**,** 9-14.

WU, K., ENG, E., KNOX, R. & CHEN, S. (2001) Demonstration of the activation of prodrug CB 1954 using human DT-diaphorase mutant Q104Y-transfected MDA-MB-231 cells and mouse xenograft model. *Arch Biochem Biophys,* 385**,** 203-8.

YAMADA, K., HARA, N., SHIBATA, T., OSAGO, H. & TSUCHIYA, M. (2006) The simultaneous measurement of nicotinamide adenine dinucleotide and related compounds by liquid chromatography/electrospray ionization tandem mass spectrometry. *Anal Biochem,* 352**,** 282-5.

YANG, H., YANG, T., BAUR, J. A., PEREZ, E., MATSUI, T., CARMONA, J. J., LAMMING, D. W., SOUZA-PINTO, N. C., BOHR, V. A., ROSENZWEIG, A., DE CABO, R., SAUVE, A. A. & SINCLAIR, D. A. (2007) Nutrient-sensitive mitochondrial NAD+ levels dictate cell survival. *Cell,* 130**,** 1095-107.

ZENNO, S., KOBORI, T., TANOKURA, M. & SAIGO, K. (1998) Conversion of NfsA, the major Escherichia coli nitroreductase, to a flavin reductase with an activity similar to that of Frp, a flavin reductase in Vibrio harveyi, by a single amino acid substitution. *J Bacteriol,* 180**,** 422-5.

ZENNO, S., KOIKE, H., KUMAR, A. N., JAYARAMAN, R., TANOKURA, M. & SAIGO, K. (1996a) Biochemical characterization of NfsA, the Escherichia coli major nitroreductase exhibiting a high amino acid sequence homology to Frp, a Vibrio harveyi flavin oxidoreductase. *J Bacteriol,* 178**,** 4508-14.

ZENNO, S., KOIKE, H., TANOKURA, M. & SAIGO, K. (1996b) Conversion of NfsB, a minor Escherichia coli nitroreductase, to a flavin reductase similar in biochemical properties to FRase I, the major flavin reductase in Vibrio fischeri, by a single amino acid substitution. *J Bacteriol,* 178**,** 4731-3.

ZENNO, S., KOIKE, H., TANOKURA, M. & SAIGO, K. (1996c) Gene cloning, purification, and characterization of NfsB, a minor oxygen-insensitive nitroreductase from Escherichia coli, similar in biochemical properties to FRase I, the major flavin reductase in Vibrio fischeri. *J Biochem,* 120**,** 736-44.